

The Public Health Impacts of Pollen on Allergic Rhinitis and Asthma

Submitted by Luke Talbot to the University of Exeter

as a thesis for the degree of

MSc by Research in Medical Studies

In September 2017

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: .....

## Abstract

Pollen grains produced by grasses, weeds, and trees are some of the most common stimuli associated with exacerbations of allergic rhinitis and asthma. The pollen seasons can be associated with significant health impacts for those affected, and evidence showing the species involved and the monitoring of pollens would prove extremely useful.

This study investigated variations in a range of vegetation types across England (as a proxy for plant-specific pollen exposures), and associations with prescription rates. The study used data from General Practices (via NHS Digital), the Met Office and other secondary data sources. Radiuses drawn up proportional to the population density of each General Practice were used to measure the vegetation densities within each one, as an estimate of the practice population exposure. Area data concerning a range of confounders and effect modifiers were run against the prescription rates of relevant medications per 1000 patients in linear regression models.

The aim of this project was to estimate associations based on this data alongside a range of other factors including air pollution, population ages, and deprivation indices. This study allows for spatial data analysis across England, and to compare the various vegetation species with the prescription rates in the months of their active seasons, adding to our understanding of the impacts of pollen on allergic disease outcomes. Key findings indicated a mixed collection of results. Some indicated positive associations between some vegetation densities and prescribing rates for respiratory and allergic conditions, but other results suggested inverse associations, contrary to hypotheses. There was also some suggestion of effect modification by air pollution, with greater adverse effects of grass density in areas with higher PM<sub>10</sub> concentrations.

## Table of Contents

<b>1 Introduction</b>	<b>6</b>
1.1 Allergic Rhinitis & Allergic Asthma	6
1.2 Pollen	8
1.3 Study Aims and Approach	9
1.4 Structure of Thesis	9
<b>2 Literature Review</b>	<b>10</b>
2.1 Pollen Exposure and Allergic Asthma/Rhinitis	11
2.2 Confounders & Effect Modifiers	15
2.3 Climate Change	19
2.4 Weather Effects	20
2.5 Summary	22
<b>3 Methodology</b>	<b>24</b>
3.1 Creating a Single Full Dataset	24
3.2 Statistical Analysis	32
<b>4 Results</b>	<b>35</b>
<b>5 Discussion</b>	<b>58</b>
5.1 Grass Vegetation	58
5.2 Ragweed and Birch	60
5.3 Other Vegetation	62

<b>5.4 Effect Modification by Air Pollution</b>	<b>65</b>
<b>5.5 Association with Confounders</b>	<b>67</b>
<b>5.6 The Study</b>	<b>69</b>
<b>6 Conclusion</b>	<b>75</b>
<b>Acknowledgements</b>	<b>77</b>
<b>Reference List</b>	<b>78</b>
<b>Appendices</b>	<b>91</b>
<b>1. Stata File Do-script</b>	<b>91</b>
<b>2. GME Commands</b>	<b>117</b>

## **Tables & Figures**

<b>Table 3.1:</b> Data sources for tree density maps produced by McInnes et al, 2017	<b>27</b>
<b>Figure 3.1:</b> Formation of the variable radiuses around General Practices in ArcGIS	<b>29</b>
<b>Figure 4.1:</b> Graphs showing the prescription rates vs. months	<b>39</b>
<b>Table 4.1:</b> Descriptive Statistics for all key variables	<b>36-37</b>
<b>Table 4.2:</b> Regression Model 1	<b>41</b>
<b>Table 4.3:</b> Regression Model 2	<b>42-43</b>
<b>Table 4.4:</b> Regression Model 3	<b>46-47</b>
<b>Table 4.5:</b> Regression Model 4	<b>49-50</b>
<b>Table 4.6:</b> Regression Model 5	<b>51</b>

<b>Table 4.7:</b> Regression Model 6	<b>53</b>
<b>Table 4.8:</b> Regression Model 7	<b>56</b>
<b>Table 5.1:</b> Table of Correlation Coefficients between pairs of GP-level variables	<b>72</b>

## **1 Introduction**

### **1.1 Allergic Rhinitis & Allergic Asthma**

The summer months can be a pleasant time, with many more people going out for a wide range of activities, from city visits to outdoor exercising. However, this time of year also coincides with the pollen seasons of many plants, and thereby to many comes as a curse in the form of disorders like asthma and allergic rhinitis, the latter known more commonly as hay fever. A common disorder seen in a significant percentage of populations in the UK and around the world, allergies have a wide range of severities, and can cost health services upwards of £900 million for primary care in the UK alone (Allergyuk.org 2016). Asthma and rhinitis have always been of great scientific interest, particularly due to the complex nature of the onset of the conditions, which occurs through a combination of environmental and genetic components as multifactorial interactions. Some exposures can affect disease onset/exacerbations more than others, with pollen in particular one of the most well-known and well-researched activation stimuli (D'Amato, Cecchi et al. 2007, Sih and Mion 2010). In more recent years, the number of those affected has steadily risen around the world, and is thought in part to be due to some confounding factors and issues such as increases in air pollution and less access to rural areas/greenspace, amongst other consequences that have amassed due to a rising world population (Pawankar, Canonica et al. 2011). Another suggestion is that of the “Hygiene Hypothesis”, which puts forward the proposal of the increasingly “clean” environments within houses. Infants and young children coming into contact with less natural environmental stimuli such as dirt is thought to lead to the prevention of a the usual immune response to this scenario, and could result in an exaggerated reaction when an individual is exposed to the stimuli (Okada, Kuhn et al. 2010). The mechanisms through which such conditions arise can be complex and have multiple pathways, and this will be detailed further below in this chapter. With cleanliness and pollution becoming more and more relevant to everyday life, there has been an increased interest in how particles like pollen grains can interact and increase their sensitivity, which could be consequential for the high number of rhinitis and asthma sufferers both in this country and around the world (D'Amato, Bergmann et al. 2014).

Allergic Rhinitis is a common disorder of the respiratory system, affecting the upper respiratory tract, with particular significance on the nasal passages. A common global disorder and with multiple variations, the condition affects up to 25% of the European population, and approximately 500 million people worldwide (Bousquet, Khaltaev et al. 2008). Allergic asthma, by comparison, is the most common form of asthma presented globally, and affects the lower respiratory tract, in this case the lower areas of the lungs; the bronchioles and alveoli, and has

a global prevalence of 300 million (Masoli, Fabian et al. 2004). Both conditions can be mediated by antibodies known as immunoglobulin E (IgE), and upon the stimulus exposure to these immunoglobulins, they become bound to specialised white blood cells known as antigen-presenting cells. Upon activation, these undergo mass degranulation; a process which results in the release of inflammatory mediators and cytokines into the bloodstream, which are targeted to the affected area(s), and initiate functions to combat the supposed foreign substances, including inflammation, vasodilation and further immune cell recruitment (Marieb and Hoehn 2013).

Both conditions affect people of all ethnicities, countries, and ages, however, infants and those with existing respiratory conditions are more much likely to find themselves affected. Although the conditions are separate, the two have been shown to interact with one another, and ultimately result in a combined disorder involving the entire respiratory system, with varying symptoms and resulting health effects (Makra, Matyasovszky et al. 2014) (Sheehan, Rangithienchai et al. 2010). There are other methods of allergen sensitisation as well. One primary mechanism for the hypothesis of its onset establishes an immune response deviation being responsible for the onset of allergies. The Th1/Th2 hypothesis focuses on the pathogen-specific (adaptive) immune response, and suggests that within the adaptive immune system, a balance exists between two types of response within adaptive immunity: the onset of inflammation (Th1 subset immune response) and inflammation mediation (Th2 subset immune response) that help with recruitment, attacking, and monitoring of the specific immune response. The two subsets keep each other balanced, to ensure an immune response is controlled. However, in the case of allergies and asthma, upon contact with certain stimuli, there can be an imbalance that favours an uncontrollable inflammatory activation, where Th1 outweighs the Th2 mediatory response. Through this, inflammation and immune cell recruitment vastly increases in an exaggerated immune response, which is seen following exposure to the stimuli, and acts as the resulting rhinitis/allergic response, presenting in the symptoms seen by those affected (Okada, Kuhn et al. 2010).

Both rhinitis and asthma can be classed into allergic and non-allergic phenotypes which are separated by the presence or absence of IgE (present in allergic phenotypes). The severity of these responses can vary wildly from mild, local responses at the point of contact; where only inflammation and mild irritation occurs, to a full-scale, severe systemic response; where systemic rashes and inflammation can be seen, and in the worst cases anaphylactic shock, which can prove fatal (Stone, Prussin et al. 2010). As a result of this, alongside high incidence rates, dual-interaction, and the potential harm that can be caused, both conditions are of high

scientific interest, and thorough research is conducted on various aspects of both rhinitis and allergic asthma almost continuously around the world.

## **1.2 Pollen**

Pollen grains are small granules released by most species of plants during reproduction. Miniscule in size, these grains contain the genetic information by which a plant can reproduce. Pollen grains leave the anther of the male parent plant through a number of means, and though some will never reach their desired destination, others will eventually come into contact with another flower of the same species, landing on the stigma of a female parent plant. The grains then initiate the growth of a pollen tube, down into the stigma of the recipient plant, and into the ovum. The genetic information is passed down into the ovum, and the reproduction process begins (Barrett 2002).

Pollen from grass, weed, and tree families is known to be responsible for high levels of allergenicity during the spring and summer months. This varies with local geography, with southern areas of the UK having earlier season starts than areas further north. The period of the season is a time of high interest for research concerning both pollens of various plant species, and changes in incidence rates of respiratory conditions. The grass family Poaceae is of particular interest. It is the most common grass family worldwide, and its pollen is an important and known risk factor for rhinitis in the UK (Christenhusz and Byng 2016, Metoffice.gov.uk 2016). It is also known to interact with a number of variables that can alter the levels of pollen in the environment and atmosphere. Grass species can have several methods of reproduction both in the presence and absence of pollen. A primary method is the dispersal of pollen either through the use of bright and colourful flowers, where it can be collected by insects and other animals, or through wind, which subsequently releases loose pollen grains into the atmosphere from plain flowers or nonflowering plants (Friedman and Barrett 2009). Poaceae species can be found naturally on every continent, and with more than 7000 genera and 12,000 species, the family makes up a substantial proportion of the base of the food chain for many thousands of animals, including humans, in ecosystems around the world (Schüler and Behling 2011). The pollen seasons begin and end differently for various plants, with alterations seen between families, genus, and species all noted. Grass pollen seasons usually start in May and end in July/August, although various conditions can alter the season to start earlier, or end later. Poaceae is a particular grass family of interest, as it is responsible for a large majority of hayfever cases, including 95% of cases in the UK (Metoffice.gov.uk 2016).



### **1.3 Study Aims and Approach**

The aim of this Masters by Research project is to investigate and establish any relationships found that could suggest an association between pollen exposure and vegetation density, and rates of prescriptions administered for allergic rhinitis and asthma in England. There is a particular focus on pollen of the Poaceae grass family, but several other key allergenic plant species will also be investigated. The approach will be to apply the methods of environmental epidemiology using large secondary datasets that include measures or indicators of the exposures and health outcomes of interest.

Using regression analysis on the prescription rates and vegetation, alongside many measurable effect modifiers and confounders, and populations based on LSOAs in England, this study will aim to achieve a better understanding of the impact of pollen (particularly Poaceae) on allergic health outcomes across England, considering the impacts of other factors such as locality, land use, and weather. The results will be placed in the context of existing based on a review of the literature on the importance of pollens and the environment on allergic rhinitis.” This information will prove useful in estimating which areas are the most affected, which will prove useful to those who find periods of the year difficult to manage as a result of high pollen levels.

### **1.4 Structure of Thesis**

Chapter 2 presents a review of the existing literature on the relationships between pollen exposure and allergic rhinitis and asthma, and related environmental risk factors. Chapter 3 describes the methods used for the study; the collection of data, the use of spatial and statistical software, and an analysis plan. Chapter 4 presents the results produced, and these are discussed in chapter 5. Chapter 6, the final chapter, brings together the findings and draws conclusions with reference to the research question, as well as raise the potential for further study on the topic.

## 2 Literature Review

To utilise existing research and the evidence shown, a literature review was undertaken to collect relevant papers that could provide useful context and help compare the findings from this study with those already collected. The aim of this literature review was to establish published literature concerning the impact of pollen exposure on respiratory conditions, particularly allergic rhinitis and asthma, and investigate the varying effects between vegetation species in order to establish the best pollen species candidates for investigation within our obtainable data. To approach and achieve this aim, the investigation of the literature addressed several areas that would enable the development of an understanding of the topic. Commencing with establishing current research and relationships between pollen exposure and asthma and rhinitis, following which the methods of collecting data and subsequent analysis of these studies will be scrutinised. Finally, this review will set out and summarise knowledge regarding other risk factors that may also play a part in how pollen can affect the consequent health outcomes.

A search criteria was devised, and databases identified for the searches to be carried out. The search criteria included terms for the subject of interest, being “Pollen”, as well as the area of interest, in this case being the variables within the outdoor environment with terms of “Outdoor”, “Urban”, and “Rural”, as well as the particular conditions of the studies, such as “Asthma”, “Sensitisation”, and “Anaphylaxis”. The following databases were used to carry out the search terms (discussed beneath):

- Scopus
- Pubmed
- Web Of Science
- Cochrane (Wiley Online Library)
- Science Direct

[“Pollen”] AND [“Environment” OR “Environmental” OR “Outdoor” OR “Rural” OR “Urban” OR “Geographical” OR “Geography”] AND [“Asthma” OR “Asthmatic” OR “Allergy” OR “Allergic” OR “Hypersensitive” OR “Hypersensitivity” OR “Rhinitis” OR “Sensitisation” OR “Anaphylaxis”].

To ensure the most recent evidence of relevance was incorporated into this review, the search criteria were applied to the publication databases, and research conducted between the years 2001 and 2016 was included with research articles and literature reviews considered drawn from studies around the world. The initial search of “Pollen” was the primary subject of interest during the initial literature search, in order to include the wide range of plant pollens

that might impact on the health outcomes of interest. However, whilst analysing the literature, it was deemed that in particular, Poaceae (grass), ragweed, and birch tree pollens were the most relevant and significant candidates fit for focused investigation, due to the high incidences of rhinitis resulting from exposure to pollen from these species. The inclusion criteria thus includes papers with mentions of the pollens of at least one of the three above species in the title or abstract, and published at any point between 2001 and 2016, given developments in the field in recent years. Applying this search and inclusion criteria to the databases, 10839 references were returned, and following the removal of inaccessible and irrelevant articles, the final reference number of relevant citations reached 1576. Whilst many papers were directly relevant to the search criteria, many instead focused on similar conditions such as dermatitis and conjunctivitis, and henceforth excluded. Many of the most relevant studies included in the final number utilised primary data in conjunction with secondary, collecting pollen using volumetric traps in areas of interest, and analysing these against datasets obtained from other sources, such as hospital episode statistics and weather forecast records. Others used solely secondary data to determine any relationships suspected to occur between pollen exposure estimates and relevant health outcomes.

## **2.1 Pollen Exposure and Allergic Asthma/Rhinitis**

Measuring atmospheric pollen concentrations continuously is an important means of determining the changing levels of pollen and pollen species in the local environment. Data is collected by the use of volumetric traps to collect pollen species samples and counts are made of the number of grains. Traps are normally located on rooftops in the areas of interest, so as to collect a significant level of pollen from a wider area, which is not necessarily found at ground level (Peel, Kennedy et al. 2014). Once the grains are collected in the trap, they are examined and recorded at regular intervals, allowing measurements for particular allergenic pollens in the area, and, with sufficient data, for forecasts to be made in the years to come (García-Mozo, Galán et al. 2009). This also acts as a reliable method to track pollen seasons, as well as the transport of grains, and can aid in identifying other factors that could affect pollen counts.

There are limitations regarding this collection of pollen, however. Some grain types can be indistinguishable from others under microscopic examination, making it hard to tell the exact species of plant from which the grains of pollen originated, grass pollens being one primary example, usually resulting in a collective count. On the whole this is not considered a major issue, as the pollens of plant families such as Ambrosia or Poaceae can often be identified, even if the species themselves cannot (Kraaijeveld, de Weger et al. 2015). Another associated

limitation is that although monitoring stations are able to detect pollen grains in the particular area, there is no way to necessarily know their exact source of origin, with factors such as weather and climate zones considered to be two significantly associated factors with pollen dispersal and spread, particularly wind, which has been shown as a major variable in the distance pollen grains can travel (Smith, Emberlin et al. 2005). These particular issues are described in sections 2.2 and 2.4. Exposure to the pollens by people with allergic asthma and/or rhinitis varies through a number of factors (see confounding factors – section 2.2), with some needing medical treatment following exposure with varying severities, from a visit to their GP for medication, to urgent care at the nearest hospital.

A number of studies have investigated pollen levels and health outcomes using secondary data in the form of hospital admission trends (e.g. Johnston and Sears 2006, García-Mozo, Galán et al. 2009, Hess, Darrow et al. 2010, Gonzalez-Barcala, Aboal-Vinas et al. 2013) . Data collected by hospitals concerning the admission of patients with asthma, rhinitis, or similar symptoms have been used in many studies to record the trends of asthma, rhinitis, and other respiratory disorders throughout the pollen season.

Emergency Department attendances and admittances have helped show how combined local events such as weather conditions and increased pollen levels can have an effect on health. The relationship between pollen and hospital attendances has been tracked and has progressively led to the formation of a pattern, showing the increase in hospital admissions during late spring and summer periods and normally coinciding with the release of pollens, and decreasing again with the temperatures dropping and the summer pollen seasons drawing to a close (Johnston and Sears 2006, García-Mozo, Galán et al. 2009). Although not all people affected need attend hospital, annual trends of patient admissions with asthma and rhinitis-like symptoms correlated with the recorded grass and ragweed pollen season times, showing a positive relationship between exposure and the health outcomes. When examining trees, trends were seen for pollens, such as pine, but was not necessarily as strong for birch pollen however, nor could it be concluded without doubt that pollens alone were responsible for these admissions (Zhong, Levin et al. 2006, Feo Brito, Mur Gimeno et al. 2007, Dales, Cakmak et al. 2008). Many hospitals around the world, including Australia and the Iberian Peninsula, where many studies have been centred, have found that during the summer months, the number of asthma admissions for those under 60 undergoes a dramatic increase, and has been associated with raised pollen levels, which could be considered surprising, given the incidence rate for rhinitis and asthma is not necessarily higher within this population (Erbas, Chang et al. 2007,

Gonzalez Barcala, Aboal Viñas et al. 2010, Mathur 2010). The social-demographic factors, including age, are discussed further below in section 2.2.1.

Whilst it would be difficult to source the exposure back to specific pollen species, it is likely the increase in admissions for respiratory illnesses throughout the summer period is in part due to pollen seasons at the time, with concentrations strengthened by the increased temperatures and daylight hours resulting in increased pollen release by plants (Bates, Baker-Anderson et al. 1990, Hong, Zhong et al. 2016). Other causes have been theorised as having significant effects on pollens, their release, and subsequent respiratory responses within individuals. Whilst researching asthma hospital trends throughout the summer months in Vancouver, Canada, a noted increase in air pollution levels in early autumn were recorded, as the grass pollen season came to an end. However, there was no noted change in admissions during this period, suggesting that any changes in air pollution during the grass pollen season had little, if any observed significant effect on asthma patients (Tarlo, Broder et al. 2001). This was thought to be in part due to increased ozone levels, resulting from volatile NO<sub>x</sub>, which can be converted to ozone as a result of increased sunlight. However, many papers have also contested Tarlo's results concerning admission, as observations have been made in a number of other studies showing air pollution significantly and directly affecting asthma exacerbations. Senenchal, Visez et al. (2015), is one such study that disagrees with the previous statement. In a review that details the relationship between pollens and air pollution thoroughly, Senenchal et al. detail findings displaying physical abnormalities in urban-sourced pollen, and increased fragility of these grains. Further findings also show variations in this damage between pollens of both different and the same species, and pollutant concentrations (Ruffin, Williams et al. 2009, Lu, Ren et al. 2014). However, the authors also describe studies that have noted no major differences, as mentioned in the previous point, as well as some of the possible reasons why (Kanter, Heller et al. 2013). With little literature focusing on this particular irregularity, the resulting conflict leans towards the prospect of further research to conclude upon the most reliable and accurate results, as well as evidence to help debunk earlier findings and what, if any, confounding factors or other issues may have had a possible effect in altering any of the collected findings.

Several studies have examined the potential roles of confounders and effect modifiers in much smaller environmental changes: one such topic of interest is investigating comparisons in air pollution levels between ground and rooftop levels, to see how the effect of urban canyons (particularly in well-built up areas) could alter readings collected. The results of such research can be used in parallel with research conducted by Peel (Peel, Kennedy et al. 2014), where pollen

level variations were examined in the same setting. Whilst Peel found higher levels of pollen concentrations at a height of ~10m (rooftops), another study (Costabile and Allegrini 2007) presented results detailing significantly higher air pollution levels at ground level for a number of pollutants ('interference' such as car pollution was not counted). The two studies suggest height and particular urban settings as another factor relating to exposure to harmful stimuli. The topics concerning confounders and effect modifiers, including air pollution, will be discussed in greater detail throughout this review.

One further point of interest is the observation of time lags, where the increased levels of pollen, or the stimuli in question, do not lead to any significant immediate health consequences, but do subsequently. One particular study on air pollution, for instance, has shown that on average approximately 3 days pass before symptoms in those affected presented, and subsequently those who have difficulty managing their conditions were admitted into hospital (Lee, Wong et al. 2006). However, by that token, naturally there are many who only suffer mild symptoms in such conditions, and as such would have no need to seek medical attention during this period. This could result in many of those affected being excluded from the data used in the study, resulting in less accurate findings. The lag between elevated pollen and consequent hospital admissions has been a topic considered of great significance concerning not just the numbers of those attending hospital and when, but pollen forecasting and monitoring as well. The importance of this is stressed, as sufficient forewarning when high pollen levels are rife could go some way towards preventing adverse health outcomes in those who are particularly sensitive. The studies examined here have found that the lag can be observed in hospital trends, and in particular through the summer months. Whilst this can vary due to a number of variables, most papers have reported hospital admissions through high pollen levels 2-3 days after the levels have been reported (Tobías, Galán et al. 2003, Lee, Wong et al. 2006).

Whilst air pollution can also affect pollen grains, some studies suggest lag between pollutant exposure to the pollen grain, the grain's exposure to the individual, and the subsequent allergic response has been shown to occur through increased pollen levels. However, correlations as a result of other factors are not as apparent (Tobías, Galán et al. 2003). Other papers have argued against this, with some, such as Lee, Wong et al. (2006), who produced a study detailing air pollutant levels recorded and the dates of hospital admission, ultimately presenting a lag of 2-4 days following exposure to the pollutant in question, with other studies further confirming this (D'Amato 2011). Although, even with these studies, lag is a topic that has not appeared as a great subject of interest in most papers, and as a result the topic remains little-researched. It is, therefore, perhaps not appropriate to draw full conclusions

from the observations of these studies until more work is done to determine typical time lags between exposure and outcome for those with rhinitis and allergies.

## **2.2 Confounders & Effect Modifiers**

A number of environmental risk factors for allergic rhinitis and asthma may act as confounding variables or effect modifiers (or both) in their relationship with pollen. Air pollution, land use, and weather are all environmental variables that have been hypothesised as associated with allergic rhinitis/asthma, and shown to have some interaction in the variations of pollen concentrations. They may also be capable of directly altering the allergenicity of pollen grains; their likelihood to burst (consequently releasing the allergen within), and in fact determine the journey of the grains from their source to destination, and how they are dispersed into the local environment, and the distance travelled from release. Several of these factors are discussed here, along with additional potential confounders.

### **2.2.1 Socio-demographic risk factors**

Whilst pollen characteristics have the potential to directly affect health outcomes for individuals with allergic rhinitis or asthma through atmospheric concentrations, many other variables can indirectly influence these same health outcomes, causing alterations not necessarily to the pollen concentrations, but to the pollen grains, the local environment, and the individuals themselves (Gleason, Bielory et al. 2014). Whilst there have been noted differences in risk factors concerning sex and ethnicity, these are slight by comparison to others, with age presenting as the risk factor showing the most significant effect on this health outcome. There is a major disposition to younger individuals: both sensitisation and incidence are noted as being the highest in young children and adolescents, and reaching their lowest rates in older individuals (Aronson 2016). Socio-economic status and occupation have also been proven as relevant rhinitis and asthma risk factors, with those working in the countryside, such as farmers, seemingly most protected, whilst those working and living in urban locations most susceptible (Ogershok, Warner et al. 2007, Elholm, Linneberg et al. 2016).

### **2.2.2 Air Pollution**

Air pollution can act as both a major confounder and effect modifier on the overall allergenicity of pollen grains, with evidence to show how particulate matter (PM) at various

concentrations and gaseous compounds can not only directly affect the health of those who are exposed to elevated concentrations by being a potential cause of asthma, but also indirectly, through their interaction with the pollen grains (Gleason, Bielory et al. 2014, Janke 2014). Chronic inflammation and asthma have both presented as a consequence of direct exposure to excess pollutants in areas of low air quality, and high levels have been shown to increase the response intensity should an individual be sensitised to pollen. In fact, through Poaceae pollen exposure to the gases nitrogen dioxide, sulphur dioxide, and ozone ( $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$  respectively), structural alterations to pollen grains can occur, ultimately causing increased fragility and likelihood of their rupture, which results in the release of the contents of the grains into the environment (D'Amato, Bergmann et al. 2014, Grunig, Marsh et al. 2014). The overall rate of these gases undertaking such a process varies with the local environment, with resulting research mapping out which pollutants are most significant to health outcomes, and where they accumulate at higher levels. Factors such as climate and land use planning, which can help designate areas producing high pollution levels, are considered of extreme significance, with urban, more industrialised areas considered more at risk than rural areas, or those with more access to greenspace (Tobías, Galán et al. 2003).

PM is made up of a mixture of solid and liquid particles, such as smoke, dust, and soot, varying in size with detection ranging from requiring an electron microscope to being identifiable with the naked eye (Marcazzan, Vaccaro et al. 2001). The particles are measured in concentrations using size as a measure of categorisation, with set diameters of  $<2.5$  ( $\text{PM}_{2.5}$ ) and  $<10\mu\text{m}$  ( $\text{PM}_{10}$ ) generally used, with the latter able to penetrate the proximal airways, and the former capable of reaching lower areas (Kim, Kabir et al. 2015). Considered the most serious form of pollution in the air of many urban settlements, it has the capability to cause serious health disorders such as acute respiratory infections, cardiovascular disease, and cancers of the trachea and lung (Cohen, Ross Anderson et al. 2005), as well as resulting in dispositions to disorders such as rhinitis and asthma (D'Amato 2011). Air pollutants such as  $\text{O}_3$  and  $\text{NO}_2$  constitute the make-up of photochemical smog, and their inhalation at high concentrations can result in increased sensitivity or reactivity, and can consequently damage the airway and the lungs, leading to the onset of respiratory conditions. Air pollution has been shown to lead to an increase in asthma exacerbation incidences, and other respiratory disorders have been observed as a direct consequence of the presence of high pollutant levels (Sousa, Martins et al. 2008, D'Amato, Bergmann et al. 2014). The abundance and severity of air pollutants varies worldwide, with some, such as ozone, accumulating in warmer climates, and others, such as  $\text{NO}_2$ , being found present in more temperate areas, such as the UK, with higher levels recorded in large cities (Vardoulakis, Solazzo et al. 2011).  $\text{SO}_2$  has also been shown as a major contributor to air



pollution, and has severe implications on respiratory dysfunction. However, unlike most other constituents of air pollution, in the absence of SO<sub>2</sub>, patients have been observed making a quick and full recovery, so its effects appear to be reversible to some degree (Gennaro, Antonello et al. 2011). Hospital admissions for asthma have been shown to be increased at high pollution levels, with studies presenting an enhanced allergic response in those with rhinitis/allergy symptoms during times of high exposure through various means, be it the local geography, weather, or another variable discussed here, all have a direct effect on the concentration of air pollution, from land use to wind speed, which will be discussed below alongside further factors. (D'Amato, Cecchi et al. 2010, D'Amato, Bergmann et al. 2014).

### **2.2.3 Climate Zones and Urban/Rural Context**

The difference between urban and rural areas has long been confirmed as a major factor in the development and severity of respiratory disorders. As previously discussed, pollution levels observed in urban environments are known to be a major risk factor in the development of asthma, and as such, gaining further understanding about how different environments affect the onset of such conditions is paramount to drawing up management strategies and solutions. Climate zones were developed to counteract the inaccuracies of classifying areas as either “rural” or “urban”. Considering the fluidity of population growth and development across the world, simply classifying areas as “urban” or “rural” is overly simplistic. A new system, known as the Local Climate Zone (LCZ) system incorporates 17 zone types, each with differences in land cover, surface structures, and human activity (Stewart and Oke 2012). The LCZ system is said to provide a standard framework for classifying land for worldwide use and determining further significant environmental factors that could affect health outcomes, from which areas could be identified to be of increased risk to particular conditions compared to others, such as what is seen with asthma and rhinitis, allowing a much higher complexity in identifying a number of factors such as greenspace and air pollution that identifying areas as rural and urban cannot deliver.

Differences in local climate zones have also shown to be associated with changing pollen levels. Comparing the variation between the extreme urban and most rural environments are considered significant due to the observable high levels of air pollution associated with urban environments, and the increased exposure to pollen grains in rural areas (Ziska, Gebhard et al. 2003). The land use and the surrounding environment, combined with the climatic records of temperature and weather may help give an idea of the levels of air pollution and aeroallergens in similar climate zones around the world. If they do, it could prove useful in forecasting levels

of the incidence of respiratory conditions, as well as prove useful for forecasting pollen levels and their significance/severity in each of the different zone types.

Greenspace has also proven to be of great significance, with rural areas benefitting although receiving a greater pollen exposure, whilst urban areas show an association with increased asthma prevalence (although findings have not always proven consistent) (Andrusaityte, Grazuleviciene et al. 2016). Areas within urban areas can be managed and designed to minimise exposure to pollens associated with respiratory difficulties, but evidence has suggested that pollen exposure in rural areas can result in lower incidences of asthma and other respiratory issues, suggesting that artificial or managed greenspace, although it can be organised to minimise or remove known allergenic species, still does not have the effect seen in rural areas, though other confounders and effect modifiers likely play a role here too. Research into greenspace has led to developments touching on a number of areas long thought unrelated, and shares many interactions, providing a network of variables that can be considered further in researching the wellbeing of those it affects.

Proximity to greenspace in coastal areas is one such variable to have a noticeable effect on incidence of rhinitis and asthma, with greenspace proximity showing a substantial decrease in local air pollution and disease prevalence, as is also seen in coastal areas. Coastal areas with high levels of greenspace are considered some of the environments at the lowest risk of developing asthma or rhinitis. However, some findings have proven inconsistent as a result of several potential other factors, including weather (such as prevailing wind) and land use, with further research still being carried out (Cariñanos and Casares-Porcel 2011, Paul J. Brochu 2011). However, industrial coastal regions are often considered to have some of the highest levels of air pollution, through both the land use, and the effects of the local environment, including weather, and effect modifiers such as sea salt (Bhugwant, Bessafi et al. 2013). As a result, although greenspace appears to play a major part in the health of those with respiratory conditions, the land use has a major aspect in determining the benefits or detriments to health outcomes of those living there, with rural coastal areas seemingly proving beneficial in comparison to urban and/or industrial coastal zones (Osthoff, Roberts et al. 2008).

One other notable mention to the topic of greenspace in this study is the importance in acknowledging the role it plays on mental health. The stressors of everyday life have been of great interest in more recent years, with more people reporting high levels of stress and mental illness than ever before (Hunt and Eisenberg 2010). Access to greenspace has been shown to improve physical health and fitness, and also to affect health outcomes of various conditions (Matsuoka and Kaplan 2008). The exposure of a greener, more rural environment to those who

suffer from rhinitis and allergies may in fact prove beneficial to those who find life in urban living difficult in times of high allergenic pollen levels, and studies have investigated this (Scott, Von Korff et al. 2007, Sayin 2013). The results have shown that alongside physical health, greenspace helps to boost mental health significantly, and if the conditions are correct, or the greenspace controlled, the risks of allergens causing health risks can be reduced. Given the significance of greenspace in boosting mood and reducing stress, there is a suggestion that greenspace could go as far as to act as a protective measure. However, it can also be counter-intuitive, given increased greenspace also results in increased pollen exposure, resulting in a rise in the risk of possible health outcomes (Dadvand, Villanueva et al. 2014, Gilbert 2016)

Research concerning greenspace and pollen exposure in an attempt to link the two to the onset of respiratory conditions has so far proven inconsistent. Some papers have found that unlike those aforementioned (where exposure leads to the protection of allergic sensitisation) exposure to greenspace or environments like farming could in fact result in a higher risk of the formation of disorders such as asthma and rhinitis (Schenker 2005, Dadvand, Villanueva et al. 2014). Whilst some other findings have proven consistent with these and others haven't, why such an occurrence could take place remains unclear, with the sensitisation process still remaining largely unknown, though multifactorial aspects through genetics and the environment are known. It is widely accepted that some plant species can be more allergenic than others, with ragweed, Poaceae, and birch considered some of the most significantly allergenic of these (D'Amato, Cecchi et al. 2007), and following research considering urban greenspace exposure, city planning of urban green areas is now being carried out with more consideration, after it was suggested those in urban environments in close proximity to greenspaces were at a higher risk of developing asthma and rhinitis in comparison to those who did not (Carinanos and Casares-Porcel 2011, Andrusaityte, Grazuleviciene et al. 2016).

## **2.3 Climate Change**

Climate change has in recent years become a thoroughly-researched topic and one for many public and scientific discussion and debates. The warming of the atmosphere through the release and trapping of greenhouse gases has had many consequences for the global population, from warmer temperatures and rising sea levels to health effects and political action (IPCC 2014). How there is a potential for climate change to be associated with increased numbers of individuals suffering from respiratory conditions is a topic of great interest, and is being thoroughly investigated and researched. A number of compounds and molecules have increased concentrations in the atmosphere through climate change, including those mentioned in

previous sub sections of this chapter. However, the effects afore-mentioned are not their only consequence, as will be discussed here.

The increases in greenhouse gases around the world have had a number of detrimental effects on both locations and their populations; many more individuals are exposed to levels of low air quality, and the number of those sensitised to stimuli and developing conditions such as rhinitis and asthma is rising steadily worldwide (von Hertzen and Haahtela 2005, Lake, Jones et al. 2017). Weather patterns have also changed, with more extreme weather recorded worldwide, from increases in temperatures to higher intensity storms, a lot of these weather examples are considered highly unusual in affected areas (Kerr 2007). These changes have been perceived as a major issue, particularly air pollution in both urban and industrial areas, where the highest levels of pollution are recorded. The higher levels, whilst harmful through a number of methods, have also been shown to affect pollen levels, and consequently affecting those suffering from pollen allergies (Monks, Granier et al. 2009). The same has been shown for weather changes, where earlier, warmer, temperatures have led to pollen seasons starting earlier, and lasting longer than expected. Worse still is that the predicted frequency of future extreme events, such as droughts, rainfall, and thunderstorms, is expected to increase exponentially over the coming decades (Barnes, Alexis et al. 2013, Todea, Suatean et al. 2013). Such changes could make forecasting future pollen levels much more difficult through previous year counts, and result in difficulties for planning ahead (Barnes, Alexis et al. 2013).

## **2.4 Weather Effects**

As mentioned above, through the monitoring, forecasting, and recording of pollen seasons of various plant families over the years, one of the primary factors involved in directly affecting atmospheric pollen concentrations has proven to be the weather. Temperature, wind speed, humidity, and precipitation are primary examples of particular meteorological factors that play roles in the process of pollen distribution, from the maturation of the plant, to the distribution of the grains in the atmosphere, and the changes in grain concentration over the season (Kasprzyk 2008). These processes are subject to the local environment, such as coastal proximity, greenspace, and urban and rural environments. Higher temperatures and longer durations of sunlight are positively associated with the release of pollen, and higher wind speeds can result in an increase of the dispersion area. The spatial movements of pollen are one such major factor subject to these meteorological conditions. Atmospheric pollens have been reported to be carried as far as 50-100km per day in the UK. One example of this is through a study by Smith and Emberlin (2005). Using volumetric traps on rooftops at Worcester University

alongside Met Office data of weather conditions in June 2001, they were able to estimate the source of some pollens in the area as those originating from more than 100km away, brought to the local environment by prevailing winds. This is just one example that confirms it is not always expected for the pollen grains present in one area to be sourced from it. However, this is again subject to many conditions, including weather, and this must be taken into account when pollen levels and spatial movements are being recorded. Whilst it is common for pollen to travel this distance in a day more extreme distances have been recorded, with some travelling as far as several thousand kilometres in the atmosphere (Smith, Emberlin et al. 2005).

On the other hand humidity and precipitation have both shown to lead to notable decreases in airborne pollen grain concentrations. Times when the environment is dominated by humidity and regular rainfall tend to have lower pollen levels in the atmosphere. Although the pollen seasons normally last between May and August in the northern hemisphere for the majority of allergenic pollens (primarily for grass and weeds. Tree pollens such as birch normally release their pollens around mid-late winter), that is not to say pollen grains are entirely absent in the winter months, as allergenic pollen grains have been recorded in the atmosphere of northern hemisphere countries from February through to October, and some events where rainfall and humidity is frequent in fact can result in a massive increase in not just pollen levels, but health-related consequences as well, which will be discussed below (Peternel, Srncic et al. 2004).

One major example of relevant significant health effects are extreme weather conditions. Some such events including high winds and storms have been shown to have a significant effect on aeroallergen levels, leading in the most extreme cases to epidemics, in particular for those suffering from asthma, rhinitis, and other respiratory conditions (Davidson, Emberlin et al. 1996, D'Amato, Holgate et al. 2015). One of the most common examples that has resulted in the recording of health effects directly associated with weather across the world are thunderstorms. During thunderstorms the conditions are ripe for pollen grains to rupture through osmotic shock, releasing the contained allergens into the local atmosphere, where they can be inhaled, penetrating into the lower airways and prompting rapid immune responses. One such well-documented scenario occurred in the UK on the 24<sup>th</sup> June 1994, where over a 30-hour period in London and the South West following a thunderstorm, 640 people were admitted to A&E departments across the region – estimated to be 10 times higher than the same rate in usual circumstances (Davidson, Emberlin et al. 1996). One point of interest was that during this episode, of those admitted, 283 had never reported any asthmatic symptoms before. This could be cause to consider the consequences of such risk factors not just for those with known

respiratory conditions. It is also worth considering that only those severely affected would have attended hospital, and many more individuals could have suffered mild effects, but self-managed their condition at home, having not felt the need to attend hospital (Venables, Allitt et al. 1997, D'Amato, Liccardi et al. 2007). Other such examples have occurred elsewhere in the UK in Birmingham, and other countries including Italy, Canada, and Australia (Packer and Ayres 1985, Girgis, Marks et al. 2000, Wardman, Stefani et al. 2002, D'Amato, Liccardi et al. 2005, D'Amato, Liccardi et al. 2007).

## **2.5 Summary**

The literature concerning the role pollen plays on respiratory conditions such as allergic rhinitis and asthma is plentiful, with the number of papers published specialising on the topic increasing each year. With that which has been published so far, a better understanding of the interactions with pollens, their environment, the immune response, and variations between species has been made, and the rising incidence of effect modifiers such as air pollution is raising a higher awareness about risks and health outcomes. However, more understanding about exposures and asthma and allergy onset must be researched, taught, and understood, as air pollution is but one major factor associated with asthma and rhinitis onset.

Of the further topics discussed above, there is also a lot of literature specialising in researching pollen distribution, land use, and air pollution. These all play major roles in pollen exposure, its interaction with the surrounding environment, and the risk of intolerance to pollen being developed, as is discussed in this chapter. However, further risk factors associated with all of these areas, including seasons and weather, appear to be under less scrutiny, and searches within these topics of interest returned only a small volume of literature associated with these, by comparison to the previous topics in the paragraph above. There is still a lot of information that could be obtained on such topics, and given the proven relationship between the weather and cases of rhinitis, the effects of the weather on pollen grains should not be underestimated.

Several risk factors mentioned in this study have been shown to correlate with hospital admission trends, and may indeed be directly responsible for the changes in admission rates observed during the pollen seasons. Both air pollution and weather effects have been shown to influence the severity of both asthma and rhinitis, resulting in higher admissions, depending on the concentrations, and subsequent weather, with mild temperatures and lessened humidity associated with such episodes (Kasprzyk 2008). Extreme weather effects, as mentioned, have also been shown to have some of the most significant consequences for those with respiratory

conditions. However, although these have been shown to relate to low to moderate changes in rates of hospital admissions, pollen levels, acting as aeroallergens, have been shown to have the highest direct effect, with increases associated with higher admittance rates, and peak levels causing high-level changes. In fact peak exposures to all factors mentioned above can cause significant, more severe responses that ultimately require medical attention (Gonzalez-Barcala, Aboal-Vinas et al. 2013).

This literature review summarises the impacts of pollen as a key risk factor for respiratory health conditions, particularly allergic asthma and rhinitis. It also described a range of environmental effect modifiers and confounders that can also play a role in these relationships and how they are measured. The literature is plentiful and well-researched, however, given the size of this topic and the sheer number of areas that can be investigated, there are still a lot of developments and progress to be made towards gaining more understanding of this topic as a whole.

### 3 Methodology

The data used in this project is all secondary and was obtained from a variety of sources with the support of my supervisor. The following files were gathered from databases and used to produce the results shown in the next chapter:

- Prescriptions issued by General Practices across England between June 2010 and December 2012, codes of the medications given (British National Formulary), the quantity prescribed, and the costs of each), as well as files containing full GP addresses (including N and E co-ordinates from British National Grid References)
- Registered populations for each GP practice collected annually for 2010/11/12
- Lower Layer Super Output Areas (LSOAs) across England in accordance with the boundaries drawn up by the 2001 national census
- Data generated by the Met Office on the spatial distribution of trees (Alder, Ash, Birch, Hazel, Oak, Pine, Willow), weeds (Dock, Mugwort, Nettle, Plantain, Ragweed) and grass for each LSOA
- Mean annual air pollution data for NO<sub>2</sub>, SO<sub>2</sub>, O<sub>3</sub>, and PM<sub>10</sub>
- The Index of Multiple Deprivation (IMD) across England
- Classification of each practice into “Urban”, “Town/Fringe”, and “Rural”

All data were integrated geographically to produce a single dataset at the general practice level, containing all the relevant information, from which the analysis was undertaken. The methodology behind how this was carried out is explained below, with the subsequent steps in further sub-sections.

#### 3.1 Creating a Single Full Dataset

##### 3.1.1 Prescribing data

The first piece of data that forms the core section of the dataset is that of the GP prescription data. Comma separated variable (.csv) files containing the counts of prescriptions given out in GPs across England for each month were available for June 2010 to December 2012, resulting in 31 .csv files holding all the relevant information needed. These data were originally obtained from the NHS Health and Social Care Information Centre (NHS Digital, 2016). This data includes the identification of each practice by number, as well as a full postal address and northing and easting coordinates (in accordance with the British National Grid). Alongside this,



data of each prescription, showing the medication according to its BNF code, and the months administered (yyyymm). The monthly .csv files were imported into Stata (version 14, StataCorp, College Station TX) and combined into a single dataset covering the study period, the steps of which will be led through in this section. Counts of prescriptions of interest were selected from the dataset based on their British National Formulary (BNF) codes, which were selected based on two groups of drugs that could be related to allergies and asthma (BNF 2011). The selected BNF codes were:

- BNF 3.1.1: Bronchodilators
- BNF 3.2: Corticosteroids
- BNF 3.3: Cromoglicate and related therapy, leukotriene receptor antagonists, and phosphodiesterase type-4 inhibitors
- BNF 3.4: Antihistamines, hyposensitisation, and allergic emergencies
- BNF 3.5: Respiratory stimulants and pulmonary surfactants
- BNF 3.8: Aromatic inhalations
- BNF 3.9.1 : Cough suppressants

These were grouped into those specifically for allergies (BNF 3.4: antihistamines) and broader, more non-specific medications that could be given out for a range of respiratory conditions, including asthma, such as  $\beta_2$  agonists and corticosteroids (All BNF codes included above – the coding for the processes carried out in Stata can be found in the Do-file in Appendix 1). Within the first process of the do-file is a loop, the function of which splits the dataset into two depending on the medications given. This ultimately separates the two datasets by BNF code, into specific medications for allergies and rhinitis (“Antihistamines”) and other respiratory medications that cover a broader range, particularly for milder symptoms (“All respiratory medications”).

Once the data of interest had been collected into these two separate datasets, the registered populations of the GP practices (also obtained from HSCIC) were merged into both. Following the addition of this data, the prescribing rates were calculated (per 1000 patients per month), and the variables that were not needed for further analysis removed. This use of the rate to consider the medications prescribed in each practice over time allows for the fact that we would expect to see higher numbers of prescriptions in practices with larger populations. The removal of unnecessary data reduces the size of the dataset, and ensured further steps were easier to undertake. The two separate datasets were then merged back into one single set of data using the Stata “mmerge” function, removing any duplicate fields in the process. A file from HSCIC containing the full address of each GP was then merged in, identifying each practice in

accordance to a unique identification code. As stated above, the practice addresses contain not only the full lines of address, but northing and easting coordinates of each in accordance with the British National Grid as well (the full method of the merging is again shown in the do-file in Appendix 1). All data collected in this project, including prescription data collected from General Practices and LSOA data has been in anonymised, aggregated form, with ethical approval neither sought nor required for this study.

### **3.1.2 Vegetation Data**

In the absence of high resolution spatial data on pollen distribution across England, we took advantage of a novel dataset on vegetation distribution, which was used instead (McInnes, Hemming et al. 2017). The study by McInnes et al produced maps of three types of vegetation: grass, weeds, and trees. Grass was collectively identified as Poaceae as a result of difficulty identifying the pollens of different species due to the similarities between them. The species of weeds to investigate were chosen as the five identified within the National Pollen and Aerobiological Unit. Finally, the trees identified to be used within this study were chosen based both on the level of allergenicity from previous studies, and on the availability of data, with all species chosen to investigate being monitored by the UK pollen network, the National Forest Industry Report (Forestry Commission), and the Trees in Towns II data. A table detailing the data sources used to map estimated tree density for the study is shown below (Table 3.1).

**Table 3.1: Data sources for tree density maps produced by McInnes et al 2017.**

Tree Species	Common Name	Broadleaf or Conifer?	Data Sources		
			UK Pollen Network	NFI	Trees in Towns II
<i>Acer pseudoplatanus</i>	Sycamore	Broadleaf	No	Yes	Yes
<b><i>Alnus</i></b>	<b>Alder</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b><i>Betula</i></b>	<b>Birch</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<i>Castanea</i>	Chestnut	Broadleaf	No	Yes	No
<b><i>Corylus vellana</i></b>	<b>Common Hazel</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<i>Crataegus</i>	Hawthorn	Broadleaf	No	Yes	Yes
<i>Eucalyptus</i>	Gum	Broadleaf	No	No	Yes
<i>Fagus</i>	Beech	Broadleaf	No	Yes	Yes
<b><i>Fraxinus</i></b>	<b>Ash</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<i>Larix</i>	Larch	Conifer	No	Yes	No
<i>Picea</i>	Spruce	Conifer	No	Yes	Yes
<b><i>Pinus</i></b>	<b>Pine</b>	<b>Conifer</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<i>Platanus</i>	Plane	Broadleaf	Yes	No	Yes
<i>Populus</i>	Poplar	Broadleaf	Yes	No	Yes
<i>Pseudotsuga</i>	Douglas-fir	Conifer	No	Yes	No
<b><i>Quercus</i></b>	<b>Oak</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<b><i>Salix</i></b>	<b>Willow</b>	<b>Broadleaf</b>	<b>Yes</b>	<b>Yes</b>	<b>Yes</b>
<i>Taxus</i>	Yew	Conifer	Yes	No	Yes
<i>Tilia</i>	Lime	Broadleaf	Yes	No	Yes
<i>Ulmus</i>	Elm	Broadleaf	Yes	No	Yes

Highlighted species were those for which maps could be produced on this project.

This vegetation data obtained for the study was sourced from a file geodatabase provided by the Met Office. This data presents 13 different vegetation species, with one grass, 5 types of weed, and 7 tree species across each LSOA, generated from 1km species maps and then aggregated to each Super Output Area. The grass and weed species have their units displayed as a percentage cover, whilst tree species are collected as a full number. As a result of this, means were used for calculating grass and weeds, whilst the sum was used for trees. The following species were included in the study:

- Grass
- Weeds
  - Dock
  - Mugwort
  - Nettle
  - Plantain
  - Ragweed
- Trees

- Alder
- Ash
- Birch
- Hazel
- Oak
- Pine
- Willow

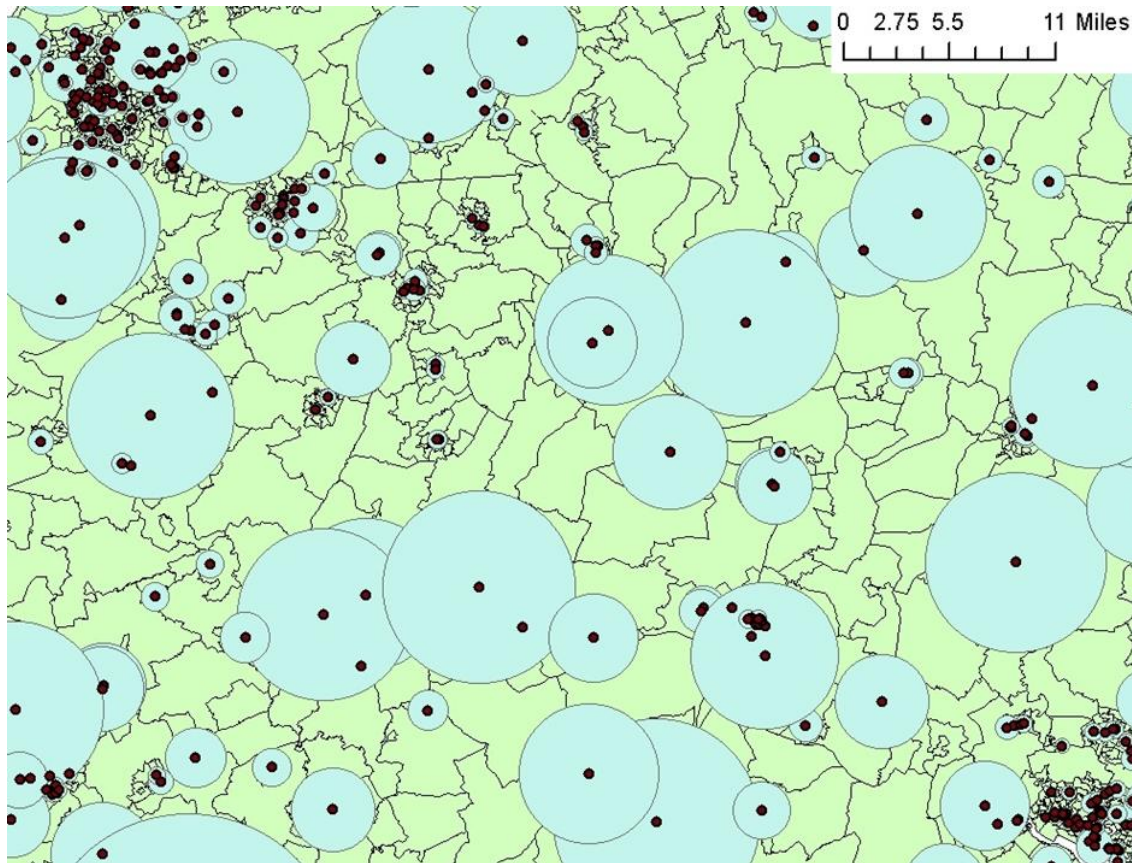
Each vegetation dataset was kept separate in 13 different Excel files. The first step was to bring these together into one single dataset, to then join to the prescription data. To start with, the means and sums of each species were identified, and all unnecessary variables in the spreadsheet removed. After this, each vegetation was imported into Stata, where they were merged together using the mmerge process. This was then exported as a .dta file, which could then be merged with the prescription data, after which, the whole dataset thus far was exported in .csv form, ready for the next step. Again, the full instructions and commands of these procedures can be found in the Do-file located in the Appendix.

Whilst the process of bringing the dataset together took place primarily in Stata, the use of a spatial software was needed in order to synthesise some of the data during the dataset's construction, particularly a method to accurately calculate the estimated vegetation in the local environment of each General Practice and its catchment. ArcGIS was chosen for its simplicity to use and the way in which it displays data, as well as the ease in importing and exporting files in various formats. As well as ArcGIS, the software Geospatial Modelling Environment (GME) was used, which will be explained in further detail below. Alongside its use in creating a single dataset, ArcGIS is also the best candidate for merging data in numerous ways to allow good visual representation, which is what it is used for with the vegetation and prescription data. Here the processes that were carried out on the data within the two softwares will be explained.

Arc GIS was first used to calibrate the coordinates in the GP prescription dataset (northing and easting) with those of the British National Grid, which were already saved into the software. The data was imported in .csv form after all separate files have been merged together and the practice addresses added, as detailed in the previous section. Upon the data's entry, a map displaying the exact location of every general practice in England was generated (see Figure 3.1). Following on from this, a shapefile containing the LSOA boundaries was imported in addition to the dataset, and joined together through a spatial join, now appearing on the map, and allocating each GP practice within its corresponding LSOA. At this point the vegetation data can now be visually shown on a map within each LSOA. However, this does not yet provide an

accurate representation of the areas surrounding the population of a GP, as individuals are not subject to a single Super Output Area, and as such this doesn't have any significance other than displaying the data per LSOA.

**Figure 3.1: Formation of the variable radius buffers around General Practices in ArcGIS**



In order to combat this, a solution was devised where a buffer of varying radius would be constructed around each GP practice in accordance with the size of its surrounding LSOA. Using this measure allows inclusion of vegetation data across more than one LSOA per practice, as well as producing larger buffer radiuses in rural areas with larger LSOAs, where the population is lower but sparser, whereas those in urban populations are far more likely to share a closer local environment. This radius' purpose is to estimate to the best accuracy the exposure to vegetation GP populations receive from their local environment. As some practices produce radiuses that encompass more than one LSOA, a weighted mean value was calculated. By using the field calculator in ArcGIS to produce a new radius value based on the shape length (perimeter) of each LSOA (m) against its area ( $m^2$ ) using the calculation below, the radius around each GP practice point was calculated relative to the LSOA it caters to, before being combined into the GP prescription dataset.

$$\text{Radius} = \left( \frac{\text{LSOA Shape area}}{\text{LSOA Shape Length}} \right) \times 2$$

The LSOA vegetation data was then calculated for each GP practice buffer as a mean or sum value. The formulation of area weighted means and sums was carried out, and as the process is complex and unable to be carried out in Arc GIS, the software GME was used (<http://www.spatialecology.com/>). This software contains a vast array of functions for statistical analysis of spatial data. For the construction of these two new variables, two functions were needed. The first, a command known as an isectpntpoly acquires data from a polygon data source and writes it to an input data source. In this case, the vegetation data from the polygon source, into the GP prescription data source. The second, an isectpolypoly, generates summary statistics for one polygon dataset based on another (see appendices for full use of these commands as they were carried out in GME – appendix 2). By combining the variable radiuses created in Arc GIS with that of the GP prescription dataset, and creating area weighted means and area weighted sums for the vegetation as appropriate, the values within the radiuses produced a more appropriate representation of exposure to vegetation than the raw means and sums within each GP practice's LSOA. Since grass and weed values for LSOAs were area densities (%) these were calculated as area weighted means (AWM); tree values were count of trees per LSOA, so these were calculated as area weighted sums (AWS).

After these functions were completed, the processes for spatial manipulation were carried out in Arc GIS. Although the AWS is calculated in GME, further calculations are still needed to produce accurate density values for tree coverage in each LSOA per square kilometre. To do this, once back in Arc GIS, the field calculator was used to create the new density variables with which to measure the tree species as accurately as possible. The universal equation used to calculate density is object divided by volume. Based on this, the equation used to produce the density measures in Arc GIS is as follows:

$$\frac{\text{Area Weighted Sum (tree count)}}{\text{Buffer Area}}$$

This equation was repeated for each tree species until each had a density field allocated, which were used as the primary measure for the analysis of the trees.

### 3.1.3 Potential confounders

As well as the prescription data and the vegetation, this study also required measures of some key confounders that have been shown/suspected to have some association with the outcome variables, some of which have already been discussed in the previous chapter. In this category fall the following pieces of potential confounder data:

- Index of Multiple Deprivation (IMD): The IMD data is the official measure of deprivation across England. Measured with each LSOA, it takes numerous factors into account, including income, employment, education, health, crime, and the living environment (Gov.uk, 2016). It has been shown to be associated with vegetation cover when compared with health, and as such any relationship between it and prescription rates would be important to adjust for (Wheeler, Lovell et al. 2015).
- Air Pollution: Air pollution data has been obtained for several air pollutants, with annual readings of each per LSOA. Following the literature review, it was decided that four would be investigated in accordance with the other variables in the study; ozone, nitrogen dioxide, sulphur dioxide, and particulate matter (PM<sub>10</sub>). These have all shown relationships with both respiratory conditions and vegetations/pollens, thus their importance here as a confounding variable is reinforced (D'Amato 2011). Data were obtained with the support of my supervisor, originally from DEFRA background air pollution concentration estimates<sup>1</sup>.
- Urban/rural classification: Data detailing the degree of urbanisation in which each practice is found was obtained with the support of my supervisor from UK Census 2011; detailing urban/rural indicators for LSOAs<sup>2</sup>. This data allocates each Practice into an urban setting, town/fringe, or a rural environment. As a confounding factor this variable is important, and will help identify any associations any variables have with the setting of each practice, in particular the vegetation and prescription rates.
- Practice Age Population: the final confounder included was obtained through the same source as the prescription data (NHS Digital), and included following findings in the review that the incidence of asthma is highest in young children and adolescents, yet still affects the older population (Aronson 2016). For each GP practice, the percentage of the population aged <15 and the percentage aged 65 and over were calculated.

These were integrated into the dataset as described in the section below.

---

<sup>1</sup> <https://uk-air.defra.gov.uk/data/pcm-data>

<sup>2</sup> <https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassifications/2011ruralurbanclassification>



### 3.1.4 Integrated dataset

Following the construction of the GP buffer data in Arc GIS with the vegetation data, the data was again converted to a text file allowing the dataset to once again be imported to Stata. With the construction of the radius and weighted means densities complete, the rest of the data was added to complete the dataset, ready for the analysis. Starting with the air pollution data, the pollutants not being used in the study were first removed from the Excel spreadsheet containing the data, after which the remaining pollutant variables were imported and merged in Stata. The environmental and IMD confounder variables were also be added to the dataset in this way, using the LSOA ID as a reference to line them up with the existing data. Finally, the ages of the populations within each GP were added into the dataset, again with the use of the mmerge function in Stata. The dataset was now complete in Stata, and the 'outsheet' function was used to create a .csv file, for use in Arc GIS for visual display.

### 3.2 Statistical Analysis

With the dataset now fully complete, analysis was undertaken to address the key research questions. An analysis plan was constructed, to best outline the hypotheses, the variables to be examined, planned methods, what the outcome variables will be. Given the number of species available for analysis and the limitation of time for the project, it was decided that whilst each vegetation species would be analysed, only one of each group would be subject to in-depth analysis. From the findings in the literature review, it was decided that grass, ragweed and birch would be the three used to represent the vegetation types, given their known allergenicity. This plan ensured the analysis was carried out methodically, addressing *a priori* hypotheses, and ensuring no unnecessary calculations were conducted. This analysis plan is presented below:

#### Background

- Asthma and Allergic Rhinitis incidences are on the rise around the world
- Pollen from various species of plant is an important risk factor for these health outcomes
- There is also evidence of relationships with a number of potential effect modifiers and confounders including air pollutants and access to greenspace.
- It has been established that pollens of some species are more allergenic than others



- While calculating and analysing the relationship between prescribing rate and vegetation distributions (as a proxy for pollen) is the primary focus, it must be acknowledged that air pollution, environmental geography and deprivation indices must also be considered

## Aims

- To identify and establish relationships between prescribing rates and vegetation density between June 2010 and December 2012

## Hypothesis

- $H_0$ : There is no association between vegetation density measures and population prescription rates of treatments administered for asthma and allergic rhinitis
- $H_1$ : There is an association between vegetation density measures and population prescription rates of treatments administered for asthma and allergic rhinitis

## Methods (Data Sources and Variables)

- Data: creation of a single dataset consisting of all necessary variables for analysis
- Population: Every registered individual of every General Practice in England
- Measures:
  - Variables Measured:

### **Outcomes**

- Prescribing rates for a) respiratory conditions and b) antihistamines

### **Exposures**

- Vegetation densities, focussing on grass, birch and ragweed.

### **Potential confounders/effect modifiers:**

- Air Pollutant levels
- Deprivation Indices
- Environmental categorisation
- Practice population age distribution

## Planned Analysis

- This will be a small area ecological analysis, investigating associations between LSOA and practice level prescription ratios and vegetation densities, alongside other measurable effect modifiers and confounders. Potential effect of seasons will also be investigated.
- Use of Ordinary Least Squares (OLS) regression analyses of vegetation densities against prescription rates per practice in the following models:
  - Model 1: Grass, ragweed, and birch, unadjusted for confounding variables
  - Model 2: Grass, ragweed, and birch, adjusted for confounding variables
  - Model 3: Grass, ragweed, and birch, adjusted, during pollen seasons
  - Model 4: Grass, ragweed, and birch, adjusted, out of pollen seasons
  - Model 5: Grass, ragweed, and birch, adjusted, investigating effect modification by PM<sub>10</sub>
  - Model 6: All vegetations, unadjusted for confounding variables
  - Model 7: All vegetations, adjusted for confounding variables

OLS Regression will be used for analysis since it is appropriate for the continuous outcome variable (prescription rates), and allows adjustment for multiple confounding variables (Kirkwood, 2008).

## Outcomes

- GP prescription rates for antihistamines and all rhinitis symptoms by practice, period, and LSOA between 06/2010 and 12/2012 alongside the AWM/densities of 13 separate types of plants. Focus will primarily be on:
- Prescription rate vs. vegetation abundance within each practice's radius, and any possible variations correlating with each species' pollen season.

Focus will be on all prescriptions administered during this time containing BNF codes of interest, with prescription rates calculated as per 1000 patients per month (pppm).

- Other outcomes shall include:
  - Descriptive table showing all the basic statistics collected
  - Regression Analyses of periods, prescriptions, and vegetation densities

## 4 Results

In this chapter, the methods described in Chapter 3 have been executed and the results of the models following processing are displayed here. Descriptive statistics regarding key outcome exposures and confounder variables are presented, and these are joined by regression analyses to investigate the associations between different vegetation densities and prescribing rates. For this chapter, focus will primarily be on the standardised coefficients and any comparisons between them than the unadjusted coefficients themselves.

The first set of data brought together in Table 4.1 displays the basic statistics produced for all the vegetation data, the confounders and effect modifiers; including air pollution, urban/rural, the General Practices, the prescription rates, and the total datapoints (n=practice-months) per month. One of the first observations to be noted are the count of datapoints per month, which show a decrease in the months between January and May. This is because data were available from June 2010 to December 2012 thus there are 2 years of data for January to May, and a full extra year of data for June to December. The count also varies between months since the number of practices providing data each month changed a little – not all practices provided the full 31 months of data. To illustrate the monthly variation in mean prescribing rates for the two groups of drugs, these are plotted in the graphs in Figure 4.1, and are mentioned in further detail below.

The means of the prescription rates across months in Table 4.1 show that when all respiratory/allergy prescriptions are considered, March, May through July, and December display notably increased prescription rates per 1000 population than other months. It must be noted, however, that when examining solely antihistamines, the inclusion criteria are narrower, and the noted differences in prescribing rates between months are also much less. The summer months show very little difference to levels at other times of the year, all but March and December, which again show raised mean values. The minimum and maximum values should also be noted in relation to the mean (Table 4.1); prescription rates for each of the groups range from zero into the thousands per 1000 populations, but the mean values themselves are all below 100 for all respiratory/allergy medications, and under 50 for antihistamines. This suggests the maximum values are potentially a result of extreme outliers, and this can be investigated further in the discussion chapter.

**Table 4.1 – Descriptive Statistics for all key variables**

Prescr. Rate by Month p'000/pm						
	Month	Mean	St. Dev.n	N	Min	Max
Prescr. Rate All (pppm)	Jan	81.76	73.12	17967	0.00	5615.39
	Feb	49.05	70.86	17952	0.00	5076.92
	Mar	90.06	92.23	17965	0.00	8461.54
	Apr	79.75	43.22	17884	0.00	1431.37
	May	88.68	48.70	17879	0.00	1725.49
	Jun	90.81	83.31	26962	0.00	7333.33
	Jul	89.40	75.32	26876	0.00	5333.33
	Aug	83.68	68.06	26818	0.00	4555.56
	Sep	84.58	72.75	26818	0.00	5433.33
	Oct	84.30	72.88	26778	0.00	4533.33
	Nov	85.90	75.10	26768	0.00	4800.00
	Dec	90.77	76.09	26737	0.00	5290.32
	Total	86.03	72.86	277404	0.00	8461.54
Prescr. Rate Antih (pppm)	Jan	40.00	37.46	17967	0.00	2846.15
	Feb	38.42	35.01	17952	0.00	2538.46
	Mar	42.52	41.83	17965	0.00	3615.39
	Apr	36.33	21.76	17884	0.00	607.84
	May	39.70	24.60	17879	0.00	823.53
	Jun	39.15	33.18	26962	0.00	2366.67
	Jul	39.56	31.58	26876	0.00	2300.00
	Aug	38.35	28.77	26818	0.00	1766.67
	Sep	40.22	35.46	26818	0.00	2766.67
	Oct	40.87	37.12	26778	0.00	2400.00
	Nov	41.97	35.97	26768	0.00	2533.33
	Dec	44.86	38.25	26737	0.00	2538.46
	Total	40.28	34.08	277404	0.00	3615.39
Grass AWM (%)	Grass	17.52	13.18	277404	0.00	87.53
Weed AWM (%)	Dock	0.43	0.27	277404	0.07	1.19
	Mugwort	0.08	0.01	277404	0.02	0.10
	Nettle	0.67	0.24	277404	0.05	1.00
	Plantain	1.34	0.49	277404	0.11	2.00
	Ragweed	0.03	0.01	277404	0.00	0.05
Tree Density (Trees/km <sup>2</sup> )	Alder	47.35	28.08	277404	0.00	260.32
	Ash	116.19	56.73	277404	0.00	553.94
	Birch	151.62	92.10	277404	0.00	860.06
	Hazel	120.96	117.27	277404	0.00	1170.01
	Oak	111.28	57.99	277404	0.00	445.20
	Pine	95.56	126.49	277404	0.00	2635.15
	Willow	65.88	30.38	277404	0.00	274.01

(p'000/pm = per thousand patients per month, AWM = area weighted mean,  
 µg/m3 = micrograms per metres cubed IMD = Index of Multiple Deprivation)

**Table 4.1 continued**

<b>IMD (2010)</b>		23.96	12.26	252659	2.86	68.47
<b>Population</b>	<b>Age 0-14</b>	17.34	4.02	252625	0.43	76.19
	<b>Age 65+</b>	15.34	5.88	252470	0.98	52.00
<b>Radius Size(m)</b>		1028.58	1314.18	277404	146.48	26935.71
<b>Air Pollution (<math>\mu\text{g}/\text{m}^3</math>)</b>	<b>Ozone</b>	5.27	1.75	277404	0.00	11.74
	<b>NO<sub>2</sub></b>	22.47	7.83	277404	0.00	56.27
	<b>PM<sub>10</sub></b>	17.48	2.53	277404	0.00	25.62
	<b>SO<sub>2</sub></b>	3.30	1.31	277404	0.00	10.88
		<b>Freq.</b>	<b>Percent</b>	<b>Cumulative</b>		
<b>GP Practices</b>	<b>Urban</b>	240837	86.82	86.82		
	<b>Town/Fringe</b>	26612	9.59	96.41		
	<b>Rural</b>	9955	3.59	100		
	<b>Total</b>	277404	100			

The next set of statistics in Table 4.1 details the variation in area coverage of the 13 types of vegetation under investigation. The percentage area coverage was used for grass and the weed species, and the density of trees per km<sup>2</sup>. As described in the methodology, these values are approximated for a 'catchment' (circular area) surrounding each GP practice, with radiuses approximately proportional to a practices' population density (so the environmental data captured for more rural practices is for a larger surrounding area). Initial observations show a large range in the cover of grass in the area surrounding each practice. The mean value is much lower however, suggesting a skewed distribution, and the standard deviation value is also noticeably high. The same range difference is also observed in several of the tree densities. Each has a zero density value as the minimum, but some have very high maximum values, in particular birch, hazel, and especially pine, which has the highest maximum density by far. By comparison alder and willow trees have the lowest maximum density of trees per km<sup>2</sup>. However, the mean values of each are more similar to each other, and although pine is still the highest and alder and willow the lowest, the values themselves are a lot less extreme. The high values of the standard deviations also mean some caution must be taken since there is a high degree of variation in these density measures. However, such variations in vegetation are not unusual, given the dynamic change in environments many GP surgeries are found in across England, from urban centres to remote rural locations. Finally, the weed values provide an interesting insight. Though at first the results themselves are much closer together, the maximum area density value doesn't go above 2%, so proportionally the differences in mean densities are similar to the differences seen with the other vegetation types. An interesting observation is that all types of

weeds bar ragweed are found in every GP neighbourhood created, albeit at very low area densities, but present nonetheless. Plantain and nettles both present as the most common by these results, closely followed by dock, and ragweed, interestingly, has the lowest values of the weeds.

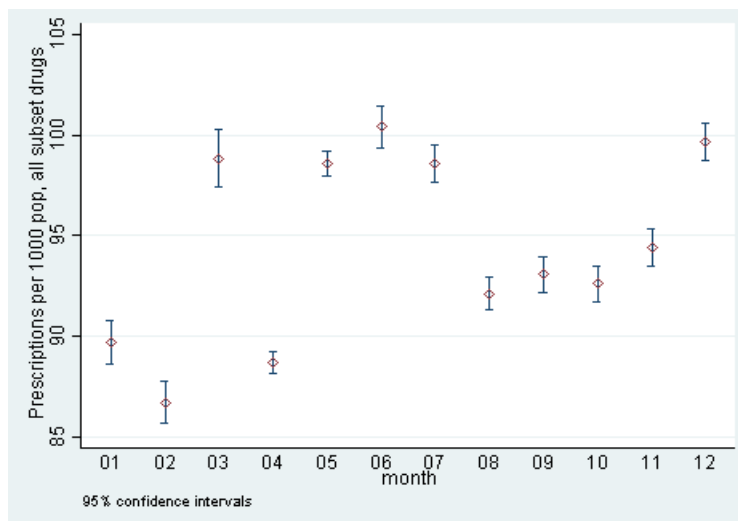
The final pieces of data that encompass the potential confounders are the deprivation indices, the population groups examined, the size variation of the radiuses, and the air pollutants measured. The IMD values range from 2.86 (least deprivation) to 68.47 (most deprivation). The data presented suggest a significant range in deprivation values throughout England. Two key population groups were examined for this project; the young (percent of GP registered population aged 0-14), who have the highest incidence rate of onset of asthma and allergies, and the elderly (percent aged 65+), who are some of the most vulnerable. These findings display that there is only a slight difference between the populations of each per GP between the mean and standard deviation values. The ranges portrayed by each show that the maximum values for 0-14 years are a lot higher than those over 65.

Of the 4 measured, the air pollution data shows nitrogen dioxide and PM<sub>10</sub> have the highest mean concentrations, whilst sulphur dioxide has the lowest. The standard deviation for particulate matter is also the lowest in accordance with its mean value, whilst the others suggest a wider variation in their values across the recorded locations.

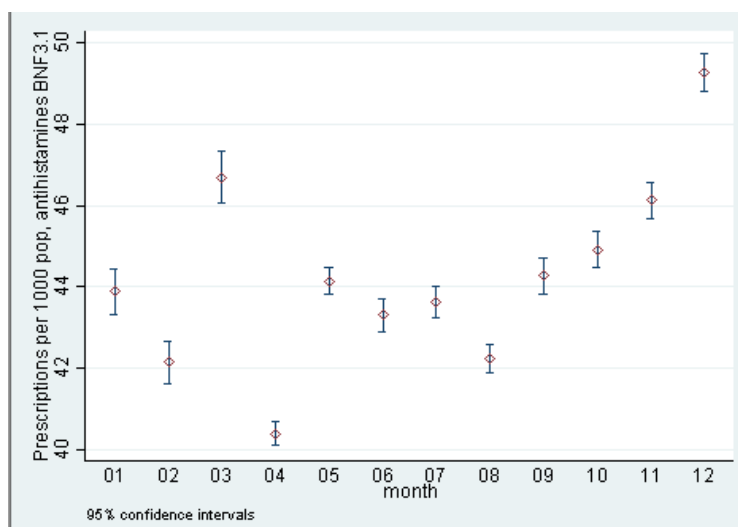
Finally, the last dataset entails the distribution of GP practices across urban-rural categories. 86.8% of General Practices fall under the definition of urban settlements, 9.6% are found in the rural-urban fringe, and the remainder in rural areas. The spread of the GP surgeries could prove interesting when considering the populations of each, and particularly the types of vegetation found surrounding each one. As described above, local population density also affect the size of the 'neighbourhood' (buffer radius) around each practice used to calculate vegetation measures.

**Figure 4.1: Graphs showing the prescription rates vs. Months: A: 'All respiratory prescriptions' (see text for definition); B: Antihistamines only.**

**Figure 4.1A**



**Figure 4.1 B**



As mentioned earlier, the graphs that compose figure 4.1 are a simple display of the two prescription rates against the mean of the collective months 2010-12 over which the data was collected. The prescription rates are naturally higher with all respiratory medications compared to just antihistamines, however, there are similarities between the two: both show high prescriptions in March and December, and low rates during January, February, and April. The rest of the months show a contrast to each other in the different figures, however. In figure 4.1 A, the summer months display high levels (May, June, and July), whilst August, September, and October show lower levels, which rise again in November, and peak in December. But when only

antihistamines are considered in 4.1 B, the summer months are considerably lower than many other months of the year. Rates are observed dropping in August, before rising up in the last few months to the December peak. Although these are basic descriptions of the data, what is shown is enough to allow discussion, which will be undertaken in the next chapter.

Following the descriptive statistics, the next set of results are from the regression models, run in a number of ways, with each examining different relationships between vegetations, prescription rates, and confounding variables. Table 4.2, shown overleaf, details the first of these regression models. This brings together the prescription rate values regressed against the vegetations of significant interest, which are grass, ragweed, and birch, as discussed in chapter 2. The values in this table (and all others detailing regression models) show the coefficient values, both unstandardised and standardised, the constant/intercept value, the 95% confidence intervals, and the p-value. The results in table 4.2 have been unadjusted for confounding variables, with adjusted values examined as part of regression model 2 (Table 4.3).

The unadjusted models indicate that GP practices in environments with higher grass densities have higher prescribing rates for both groups of medications, whilst those in areas with higher densities of ragweed and birch trees appear to be lower. The p-values are negligible at  $<0.001$ , suggesting that these associations are not down to chance. Standardised coefficients indicate the change in prescribing rates in standard deviations with a 1 standard deviation increase in the vegetation density. These show that while there is large variation in unstandardised coefficients between vegetation types, this is simply due to differences in scale between the vegetation types.

In comparison, Table 4.3 showing regression model 2 shows the regression model results for each of the three vegetation types against the prescribing rates. Again the same value types are produced, but this time adjusted for the confounding variables.



**Table 4.2: Regression Model 1: Unadjusted regression of each of the three vegetation densities against rates of the two prescription groups**

	Predictor	Coefficient	Coefficient (Std.)	95% Co		P-value
All Resp. Meds	Grass	0.147	0.027	0.126	0.167	<0.001
	Ragweed	-331.168	-0.064	-350.337	-312.000	<0.001
Antihistamines	Birch	-0.022	-0.028	-0.025	-0.019	<0.001
	Grass	0.054	0.021	0.044	0.064	<0.001
	Ragweed	-124.860	-0.052	-133.832	-115.887	<0.001
	Birch	-0.009	-0.024	-0.010	-0.007	<0.001

**Table 4.3: Regression Model 2: Adjusted regression of each of the three vegetation densities against rates of the two prescription groups**

All Resp. Meds						
Predictor			Coefficient	Coefficient (Std.)	95% Co	P-value
Grass	Grass		-0.033	-0.009	-0.044	-0.022 <0.001
	Ozone		-1.086	-0.054	-1.229	-0.944 <0.001
	NO2		-1.414	-0.312	-1.476	-1.351 <0.001
	SO2		6.176	0.217	6.075	6.276 <0.001
	PM10		-1.151	-0.079	-1.439	-1.130 <0.001
	IMD		1.590	0.524	1.578	1.602 <0.001
	Urban/Rural	Town/Fringe	6.842	0.056	6.405	7.279 <0.001
		Rural	3.900	0.020	3.243	4.558 <0.001
	Popul.n	0-14	0.780	0.083	0.747	0.814 <0.001
		65+	1.797	0.284	1.770	1.824 <0.001
Constant			52.475		50.847	54.103 <0.001
Ragweed	Ragweed		-374.121	-0.149	-389.422	-358.819 <0.001
	Ozone		-0.764	-0.037	-0.906	-0.622 <0.001
	NO2		-0.983	-0.209	-1.048	-0.919 <0.001
	SO2		5.832	0.205	5.731	5.933 <0.001
	PM10		-1.532	-0.102	-1.686	-1.379 <0.001
	IMD		1.681	0.555	1.668	1.693 <0.001
	Urban/Rural	Town/Fringe	-0.236	-0.002	-0.742	0.269 0.360
		Rural	-4.554	-0.023	-5.274	-3.833 <0.001
	Popul.n	0-14	0.717	0.077	0.684	0.751 <0.001
		65+	1.790	0.282	1.763	1.818 <0.001
Constant			57.668		56.133	59.202 <0.001
Birch	Birch		0.002	0.006	0.001	0.004 <0.001
	Ozone		-1.096	-0.055	-1.240	-0.953 <0.001
	NO2		-1.435	-0.317	-1.499	-1.371 <0.001
	SO2		6.178	0.217	6.077	6.278 <0.001
	PM10		-1.159	-0.071	-1.318	-1.000 <0.001
	IMD		1.605	0.528	1.593	1.617 <0.001
	Urban/Rural	Town/Fringe	6.556	0.054	6.131	6.981 <0.001
		Rural	3.568	0.018	2.923	4.212 <0.001
	Popul.n	0-14	0.762	0.081	0.729	0.796 <0.001
		65+	1.796	0.283	1.769	1.823 <0.001
Constant			49.860		48.278	51.442 <0.001

Table 4.3 Continued

Antihistamines						
Predictor			Coefficient	Coefficient (Std.)	95% Co	P-value
Grass	Grass		0.007	0.008	0.001	0.013 <0.001
	Ozone		-0.371	-0.035	-0.445	-0.297 <0.001
	NO2		-0.654	-0.267	-0.686	-0.622 <0.001
	SO2		3.782	0.245	3.730	3.834 <0.001
	PM10		-1.203	-0.143	-1.283	-1.123 <0.001
	IMD		0.941	0.571	0.934	0.947 <0.001
	Urban/Rural	Town/Fringe	1.090	0.016	0.864	1.317 <0.001
		Rural	-2.679	-0.025	-3.019	-2.339 <0.001
	Popul.n	0-14	0.107	0.021	0.090	0.125 <0.001
		65+	0.932	0.271	0.918	0.946 <0.001
	Constant		29.670		28.828	30.513 <0.001
Ragweed	Ragweed		-225.916	-0.164	-233.819	-218.014 <0.001
	Ozone		-0.197	-0.018	-0.270	-0.123 <0.001
	NO2		-0.400	-0.157	-0.433	-0.367 <0.001
	SO2		3.566	0.231	3.513	3.618 <0.001
	PM10		-1.394	-0.174	-1.473	-1.314 <0.001
	IMD		0.988	0.601	0.982	0.994 <0.001
	Urban/Rural	Town/Fringe	-2.932	-0.044	-3.193	-2.671 <0.001
		Rural	-7.457	-0.070	-7.829	-7.085 <0.001
	Popul.n	0-14	0.077	0.015	0.060	0.095 <0.001
		65+	0.929	0.270	0.915	0.943 <0.001
	Constant		34.256		33.464	35.049 <0.001
Birch	Birch		0.004	0.018	0.003	0.005 <0.001
	Ozone		-0.433	-0.041	-0.507	-0.359 <0.001
	NO2		-0.701	-0.286	-0.734	-0.668 <0.001
	SO2		3.766	0.244	3.714	3.817 <0.001
	PM10		-1.091	-0.129	-1.173	-1.008 <0.001
	IMD		0.948	0.575	0.942	0.955 <0.001
	Urban/Rural	Town/Fringe	1.191	0.018	0.971	1.411 <0.001
		Rural	-2.483	-0.023	-2.817	-2.150 <0.001
	Popul.n	0-14	0.097	0.019	0.079	0.114 <0.001
		65+	0.932	0.271	0.917	0.946 <0.001
	Constant		28.667		27.848	29.485 <0.001

Table 4.3 contains a large amount of information compared to the first model, and therefore much of what the data shows will be discussed in Chapter 5 alongside the rest of the models. Looking at these regressions, the associations become more complex with more variables included. The coefficient between grass and all respiratory prescriptions, now adjusted for confounding variables, is negative, suggesting an association with lower prescribing rates in areas with more grass coverage. The grass coefficient for antihistamines remains positive after adjustment. Ragweed also remains negative following adjustment for both prescription rates, and also displays the most negative values when all variables have been adjusted for, with standardised coefficient values  $< -0.1$  for both prescription rates. The regressions for birch trees flip values after adjustment, with both rates now above zero, but only 0.006 for the all medications rate, and 0.018 for antihistamines. Interestingly enough, the regressions carried out for the antihistamine prescription rates have the largest magnitude for both ragweed and birch, whilst both standardised values for grass are similar in difference, at  $-0.009$  for all medications, and  $+0.008$  for antihistamines.

The results generally indicate that in areas of higher ozone,  $\text{NO}_2$ , and  $\text{PM}_{10}$  concentrations, prescribing rates for the two drug groups are lower. Given existing evidence for these air pollutants increasing the risk of respiratory conditions, the results are opposite to what might be expected. Prescribing rates are higher in areas of higher  $\text{SO}_2$  concentrations, which is more consistent with what might be expected. Both population age groups investigated also return positive coefficients, with the highest amongst over 65s, suggesting higher prescribing rates in practices where there are higher populations within these age groups. The urban/rural variable however, shows mixed results. With both town/fringe and rural being shown as values compared against urban GPs, grass and birch models display positive values for respiratory medications, with ragweed models negative, but coefficients for rural turn negative compared to urban with all vegetations in this model when antihistamine rates are investigated. Examining the urban/rural variable using the model above, it can be shown that the values collected are quite varied with the differing vegetation. This is particularly interesting, and although grass and birch show quite strong links with the variable, this needs to be followed up in the discussion to explore reasons as to why the values vary so much. The final variable to mention, the Index of Multiple Deprivation, repeatedly returns a highly positive coefficient regression (higher with antihistamines), and at a standardised values of  $+0.5$  to  $0.6$ , making it the biggest value difference of any confounding variable observed. This indicates that prescribing rates in both drug groups are higher in more socio-economically deprived areas. These results will be discussed in further detail in the discussion chapter, in which suggestions as to why some of

these associations can be seen will be approached. All p-values fall below 0.001, confirming a statistically significant association between the variables.

Regression Model 3 follows the same equation as was used for model 2, but only uses data for the months in which each type of vegetation typically releases its pollen. The results of these regression models are presented in table 4.4 below. The months included for each vegetation's pollen season were decided from findings in the literature review (chapter 2). The months in which the grass pollen season occurs were determined as between the months of April and August, ragweed between February and September, and birch February to May. The regressions were again adjusted for confounders, and the results presented in the same way as table 4.3.

We can see from these results that as discussed with regression model 2, the same trends remain visible when only using the data within each species' pollen season. With each of the models, there is very little difference between the coefficients for the prescription rates, particularly for the antihistamine medications. When examining all respiratory medications, the difference is slightly greater, with each of the vegetations showing values further from zero during the pollen seasons than over the course of the years in general. The striking resemblance of the results in the two models when compared will make for some interesting discussion, and it will also be of interest to see how model 3 compares to 4, where only the months in which the vegetations are inactive in pollen release are included. One of the final points to note with model 3 are the p-values, in particular those for grass against the prescription rate in the antihistamine regressions, and birch for all prescriptions. Both are higher than 0.05; birch only slightly, but the antihistamine vs grass p-value is 0.261, much higher than the 0.05 conventional threshold for 'statistical significance'.

**Table 4.4: Regression Model 3: Adjusted regression of each of the three vegetation densities during the months of active pollen seasons.**

All Resp. Medications							
Predictor			Coefficient	Coefficient (Std.)	95% Co		P-value
Grass	Grass		-0.034	-0.013	-0.051	-0.018	<0.001
	Ozone		-1.445	-0.070	-1.658	-1.232	<0.001
	NO2		-1.489	-0.324	-1.582	-1.396	<0.001
	SO2		6.150	0.223	6.000	6.300	<0.001
	PM10		-0.849	-0.060	-1.080	-0.617	<0.001
	IMD		1.496	0.508	1.478	1.515	<0.001
	Urban/Rural	Town/Fringe	8.045	0.068	7.390	8.699	<0.001
		Rural	5.620	0.029	4.637	6.604	<0.001
	Popul.n	0-14	0.775	0.086	0.725	0.825	<0.001
		65+	1.908	0.309	1.867	1.948	<0.001
Constant			50.169		47.734	52.604	<0.001
Ragweed	Ragweed		-387.807	-0.152	-406.718	-368.896	<0.001
	Ozone		-0.951	-0.046	-1.126	-0.775	<0.001
	NO2		-0.985	-0.213	-1.065	-0.906	<0.001
	SO2		5.760	0.208	5.636	5.885	<0.001
	PM10		-1.365	-0.096	-1.555	-1.175	<0.001
	IMD		1.635	0.552	1.619	1.650	<0.001
	Urban/Rural	Town/Fringe	0.162	0.001	-0.463	0.787	0.611
		Rural	-3.916	-0.020	-4.806	-3.025	<0.001
	Popul.n	0-14	0.701	0.077	0.660	0.742	<0.001
		65+	1.859	0.300	1.825	1.893	<0.001
Constant			56.895		54.999	58.792	<0.001
Birch	Birch		0.004	0.009	0.001	0.006	0.006
	Ozone		-0.906	-0.043	-1.188	-0.623	<0.001
	NO2		-1.318	-0.282	-1.444	-1.192	<0.001
	SO2		5.930	0.212	5.733	6.128	<0.001
	PM10		-1.147	-0.079	-1.461	-0.834	<0.001
	IMD		1.547	0.518	1.523	1.572	<0.001
	Urban/Rural	Town/Fringe	6.429	0.053	5.592	7.267	<0.001
		Rural	3.003	0.015	1.732	4.273	<0.001
	Popul.n	0-14	0.772	0.084	0.705	0.838	<0.001
		65+	1.826	0.291	1.772	1.880	<0.001
Constant			46.060		42.945	49.174	<0.001

Table 4.4 Continued

Antihistamines							
Predictor			Coefficient	Coefficient (Std.)	95% Co		P-value
Grass	Grass		0.005	0.003	-0.004	0.013	0.261
	Ozone		-0.463	-0.042	-0.569	-0.357	<0.001
	NO2		-0.630	-0.258	-0.676	-0.584	<0.001
	SO2		3.678	0.252	3.604	3.752	<0.001
	PM10		-1.114	-0.147	-1.229	-0.999	<0.001
	IMD		0.901	0.577	0.892	0.910	<0.001
	Urban/Rural	Town/Fringe	1.427	0.023	1.102	1.752	<0.001
		Rural	-2.288	-0.023	-2.776	-1.799	<0.001
	Popul.n	0-14	0.063	0.013	0.038	0.088	<0.001
		65+	0.995	0.304	0.975	1.016	<0.001
	Constant		27.637		26.429	28.846	<0.001
Ragweed	Ragweed		-225.894	-0.165	-235.448	-216.340	<0.001
	Ozone		-0.256	-0.023	-0.345	-0.167	<0.001
	NO2		-0.372	-0.150	-0.412	-0.332	<0.001
	SO2		3.480	0.234	3.417	3.542	<0.001
	PM10		-1.372	-0.179	-1.468	-1.276	<0.001
	IMD		0.968	0.609	0.960	0.975	<0.001
	Urban/Rural	Town/Fringe	-2.747	-0.043	-3.062	-2.431	<0.001
		Rural	-7.220	-0.070	-7.670	-6.770	<0.001
	Popul.n	0-14	0.053	0.011	0.032	0.074	<0.001
		65+	0.967	0.291	0.950	0.984	<0.001
	Constant		33.160		32.202	34.118	<0.001
Birch	Birch		0.004	0.019	0.003	0.005	<0.001
	Ozone		-0.382	-0.034	-0.527	-0.237	<0.001
	NO2		-0.673	-0.265	-0.738	-0.608	<0.001
	SO2		3.686	0.243	3.584	3.787	<0.001
	PM10		-1.050	-0.134	-1.211	-0.889	<0.001
	IMD		0.927	0.571	0.914	0.939	<0.001
	Urban/Rural	Town/Fringe	1.175	0.018	0.745	1.605	<0.001
		Rural	-2.668	-0.025	-3.320	-2.015	<0.001
	Popul.n	0-14	0.101	0.020	0.067	0.135	<0.001
		65+	0.943	0.277	0.915	0.970	<0.001
	Constant		26.589		24.989	28.189	<0.001

Regression Model 4 was run as a polar opposite of Model 3. This regression uses the same outcome and explanatory variables seen in model 3, but instead of taking into account the months in which the pollen seasons take place, this model takes into account the months in which the pollens of the vegetations being investigated normally cannot be found. The months included for this model and the respective vegetation for each are as follows: Grass, from September to March, ragweed November to March, and birch, from June until January. Table 4.5, which shows these results, is displayed below.

As mentioned previously, these resulted were anticipated for comparing with those collected in model 3, during the pollen season. However, the most surprising thing to notice

about the results observed in regression 4 is the fact that the coefficients are nearly unchanged between model 4 and those observed in the previous regression. The coefficients themselves follow the same positive and negative patterns as shown previously, and again, the values for ragweed are the most significant standardised coefficients out of the variables measured. Whilst this suggests very little change between the two models, these results will be discussed in full, alongside the findings of the other regressions in the discussions chapter. Interestingly, even all but one of the p-values above the 0.05 threshold is mirrored in the previous regression model as well, with the coefficient for ozone in the antihistamine ragweed regression being the only significant p-value not mirrored in another model.

The deprivation indices maintain the most positive values, with 0.59 being the highest coefficient, again for birch/antihistamine regression. The regressions run with just the antihistamine prescriptions have the highest regression coefficients, with those containing all relevant prescriptions consistently lower. This will be discussed further in the next chapter.



**Table 4.5 – Regression Model 4: Adjusted regression of each of the three vegetation densities outside the months of active pollen seasons.**

All Resp. Meds							
Predictor			Coefficient	Coefficient (Std.)	95% Co		P-value
Grass	P.rate		-0.032	-0.011	-0.047	-0.017	<0.001
	Ozone		-0.826	-0.038	-1.017	-0.636	<0.001
	NO2		-1.359	-0.281	-1.442	-1.276	<0.001
	SO2		6.194	0.214	6.060	6.328	<0.001
	PM10		-1.600	-0.107	-1.807	-1.392	<0.001
	IMD		1.658	0.536	1.641	1.674	<0.001
	Urban/Rural	Town/Fringe	5.973	0.048	5.388	6.559	<0.001
		Rural	2.660	0.013	1.780	3.539	<0.001
	Popul.n	0-14	0.784	0.082	0.739	0.829	<0.001
		65+	1.717	0.265	1.680	1.753	<0.001
	Constant		54.145		51.967	56.324	<0.001
Ragweed	P.rate		-352.494	-0.130	-378.259	-326.728	<0.001
	Ozone		-0.468	-0.021	-0.707	-0.228	<0.001
	NO2		-0.979	-0.199	-1.088	-0.871	<0.001
	SO2		5.946	0.202	5.776	6.116	<0.001
	PM10		-1.798	-0.118	-2.057	-1.539	<0.001
	IMD		1.754	0.558	1.733	1.775	<0.001
	Urban/Rural	Town/Fringe	-0.867	-0.007	-1.718	-0.015	0.046
		Rural	-5.563	-0.027	-6.776	-4.350	<0.001
	Popul.n	0-14	0.743	0.077	0.687	0.800	<0.001
		65+	1.682	0.255	1.636	1.728	<0.001
	Constant		58.896		56.312	61.480	<0.001
Birch	P.rate		0.002	0.005	0.000	0.003	0.009
	Ozone		-1.162	-0.054	-1.329	-0.996	<0.001
	NO2		-1.475	-0.310	-1.549	-1.401	<0.001
	SO2		6.265	0.220	6.148	6.381	<0.001
	PM10		-1.164	-0.079	-1.348	-0.979	<0.001
	IMD		1.625	0.534	1.611	1.639	<0.001
	Urban/Rural	Town/Fringe	6.601	0.054	6.108	7.094	<0.001
		Rural	3.764	0.019	3.017	4.512	<0.001
	Popul.n	0-14	0.759	0.081	0.720	0.798	<0.001
		65+	1.785	0.280	1.754	1.817	<0.001
	Constant		51.187		49.352	53.021	<0.001

Table 4.5 continued

Antihistamines							
Predictor			Coefficient	Coefficient (Std.)	95% Co		P-value
Grass	P.rate		0.009	0.006	0.001	0.017	0.027
	Ozone		-0.304	-0.026	-0.405	-0.204	<0.001
	NO2		-0.671	-0.253	-0.715	-0.627	<0.001
	SO2		3.858	0.243	3.787	3.929	<0.001
	PM10		-1.268	-0.155	-1.377	-1.158	<0.001
	IMD		0.969	0.571	0.960	0.978	<0.001
	Urban/Rural	Town/Fringe	0.847	0.012	0.538	1.156	<0.001
		Rural	-2.961	-0.027	-3.426	-2.497	<0.001
	Popul.n	0-14	0.139	0.027	0.115	0.163	<0.001
		65+	0.886	0.249	0.867	0.905	<0.001
Constant			31.139		29.989	32.289	<0.001
Ragweed	P.rate		-225.865	-0.152	-239.447	-212.283	<0.001
	Ozone		-0.103	-0.008	-0.229	0.023	0.111
	NO2		-0.444	-0.165	-0.501	-0.387	<0.001
	SO2		3.703	0.230	3.614	3.793	<0.001
	PM10		-1.430	-0.172	-1.566	-1.293	<0.001
	IMD		1.020	0.592	1.009	1.031	<0.001
	Urban/Rural	Town/Fringe	-3.223	-0.046	-3.672	-2.774	<0.001
		Rural	-7.828	-0.070	-8.468	-7.189	<0.001
	Popul.n	0-14	0.117	0.022	0.087	0.146	<0.001
		65+	0.868	0.240	0.844	0.892	<0.001
Constant			35.995		34.633	37.358	<0.001
Birch	P.rate		0.004	0.018	0.003	0.005	<0.001
	Ozone		-0.451	-0.039	-0.537	-0.365	<0.001
	NO2		-0.711	-0.275	-0.749	-0.672	<0.001
	SO2		3.794	0.246	3.734	3.854	<0.001
	PM10		-1.106	-0.139	-1.201	-1.010	<0.001
	IMD		0.956	0.578	0.948	0.963	<0.001
	Urban/Rural	Town/Fringe	1.197	0.018	0.942	1.453	<0.001
		Rural	-2.419	-0.023	-2.807	-2.032	<0.001
	Popul.n	0-14	0.095	0.019	0.075	0.116	<0.001
		65+	0.928	0.268	0.911	0.944	<0.001
Constant			29.392		28.441	30.343	<0.001

The next regression model analyses the interaction of the vegetation densities with one of the potential effect modifiers considered most significant – particulate matter. The particles are theorised to interact with the pollen grains, and weaken the outer shell, making the release of the allergens within easier (D'Amato, Bergmann et al. 2014). The fifth model examines the concentration of particulate matter, and how variations in this can affect prescription rates within the LSOAs alongside the vegetation cover of the three plant species under scrutiny. As discussed in the methodology, two regressions were run in this model, one with and one without an interaction term between PM<sub>10</sub> and the vegetation density, which were compared using a likelihood ratio test. Where the interaction test indicated evidence of effect modification, models were run stratified by PM<sub>10</sub> quintile.

**Table 4.6 – Regression Model 5: All respiratory medications versus each of the vegetation types, stratified by PM<sub>10</sub> quintile.**

All Resp. Medications						
	Quintile	Coefficient	Coefficient (Std.)	95% Conf.		P value
Grass Awm	1	0.017	0.010	-0.008	0.042	0.181
	2	0.063	0.022	0.029	0.098	<0.01
	3	-0.034	-0.009	-0.078	0.011	0.140
	4	0.234	0.055	0.184	0.284	<0.01
	5	0.180	0.040	0.123	0.237	<0.01
Ragweed Awm	1	-439.772	-0.200	-485.229	-394.315	<0.01
	2	-586.908	-0.231	-635.849	-537.967	<0.01
	3	-290.078	-0.107	-340.286	-239.870	<0.01
	4	-437.338	-0.124	-485.525	-389.152	<0.01
	5	-288.350	-0.047	-370.663	-206.036	<0.01
Birch Density	1	0.013	0.044	0.010	0.017	<0.01
	2	-0.004	-0.012	-0.007	0.000	0.031
	3	-0.015	-0.044	-0.019	-0.011	<0.01
	4	-0.008	-0.017	-0.013	-0.003	0.002
	5	-0.015	-0.021	-0.023	-0.006	<0.01
Antihistamines						
	Quintile	Coefficient	Coefficient (Std.)	95% Conf.		P value
Grass Awm	1	-0.005	-0.004	-0.018	0.009	0.504
	2	0.089	0.054	0.071	0.106	<0.01
	3	0.052	0.027	0.031	0.074	<0.01
	4	0.167	0.074	0.142	0.192	<0.01
	5	0.141	0.068	0.117	0.166	<0.01
Ragweed Awm	1	-233.634	-0.186	-257.640	-209.629	<0.01
	2	-323.528	-0.228	-348.100	-298.956	<0.01
	3	-160.580	-0.115	-184.609	-136.551	<0.01
	4	-288.048	-0.155	-311.903	-264.192	<0.01
	5	-242.215	-0.088	-277.465	-206.966	<0.01
Birch Density	1	0.007	0.040	0.005	0.009	<0.01
	2	0.001	0.007	0.000	0.003	0.163
	3	-0.002	-0.012	-0.004	0.000	0.043
	4	-0.001	-0.005	-0.004	0.001	0.314
	5	0.000	-0.001	-0.004	0.003	0.879

The likelihood ratio tests for the prescription rates and particulate matter returned a chi squared p-value of <0.001, suggesting that there is a significant difference between both models in each instance the test was carried out (i.e. there is evidence of effect modification by PM<sub>10</sub> in each case). Table 4.6 displays the coefficients of regressions run stratified by PM<sub>10</sub> quintiles.

The p-values of the birch and grass coefficients interestingly, show a suggestion of statistical insignificance, with several values from measurements with each of the vegetations above the 0.05 threshold. Grass density associations suggest a stronger, positive relationship with prescribing rates of both drug groups in areas with the higher PM<sub>10</sub> concentrations. For ragweed, the apparent association with both drug group prescribing rates is inverse, and stronger at lower PM<sub>10</sub> concentrations. Finally, for birch, any pattern of association is not clear, but suggests there might be higher prescribing rates suggested with greater birch density at lower PM<sub>10</sub> concentrations, but an inverse association at higher PM<sub>10</sub> concentrations. Although these two tables show a mixture of results, they also generate a lot in the way of discussion, which can be found in the following chapter.

The previous models illustrate regressions run containing just the three vegetation types of interest. The final two exploratory models presented in this chapter include all 13 types of vegetation data obtained for this project in a single model with each outcome variable. Regression model 6 is presented in Table 4.7; this was run using the same method and outcomes as regression model 1, again not accounting for confounding variables or effect modifiers, with the difference in this case being the prescription rate compared to the increased density of all plants/trees.

**Table 4.7 – Regression Model 6: All plant/tree densities mutually adjusted, but unadjusted for confounders.**

Respiratory Medications						
	Predictor	Coefficient	Coefficient (Std.)	95% Conf.		P value
Weeds	Grass	0.147	0.027	0.126	0.168	<0.001
	Dock	16.070	0.059	15.054	17.085	<0.001
	Mugwort	-360.969	-0.059	-383.701	-338.236	<0.001
	Nettle	-19.072	-0.063	-20.192	-17.952	<0.001
	Plantain	-8.816	-0.059	-9.367	-8.265	<0.001
Trees	Ragweed	-331.168	0.064	-350.337	-311.999	<0.001
	Alder	0.083	0.032	0.074	0.093	<0.001
	Ash	-0.062	-0.048	-0.066	-0.057	<0.001
	Birch	-0.022	-0.028	-0.025	-0.019	<0.001
	Hazel	-0.031	-0.049	-0.033	-0.028	<0.001
	Oak	-0.013	-0.010	-0.017	-0.008	<0.001
	Pine	-0.003	-0.005	-0.005	-0.001	0.011
	Willow	-0.071	-0.030	-0.080	-0.062	<0.001
Antihistamines						
	Predictor	Coefficient	Coefficient (Std.)	95% Conf.		P value
Weeds	Grass	0.054	0.021	0.045	0.064	<0.001
	Dock	5.605	0.044	5.130	6.080	<0.001
	Mugwort	-157.481	-0.055	-168.117	-146.845	<0.001
	Nettle	-7.320	-0.052	-7.844	-6.796	<0.001
	Plantain	-3.117	-0.045	-3.374	-2.859	<0.001
Trees	Ragweed	-124.860	0.052	-133.832	-115.887	<0.001
	Alder	0.059	0.049	0.054	0.063	<0.001
	Ash	-0.034	-0.057	-0.037	-0.032	<0.001
	Birch	-0.009	-0.024	-0.010	-0.007	<0.001
	Hazel	-0.018	-0.064	-0.020	-0.017	<0.001
	Oak	-0.005	-0.009	-0.007	-0.003	<0.001
	Pine	-0.004	-0.017	-0.006	-0.003	<0.001
	Willow	-0.031	-0.028	-0.035	-0.027	<0.001

As can be seen here, all standardised coefficients are within a small range between +0.06 and -0.06. All but one tree species and the majority of weed regressions suggest an inverse association, with only grass, dock, ragweed, and alder producing positive coefficients for both regression sets. It is also important here to observe the significance of the vegetation not just observed in previous regressions, but those that have not been examined yet. Especially those species that show initial values that display a greater difference in regression values than the 3 that have been scrutinised. As is observed, mugwort has a larger regression coefficient than ragweed, and the regression coefficients of alder, ash, hazel, and willow are larger than that seen with birch. This will be detailed further in the final regression. However, these results do show that as with ragweed, the difference in values with mugwort is also small when presenting the unaltered coefficients with the standardised ones. The tree density associations, however, remain nearly unchanged by comparison.

Regression model 7 is the final and by far the largest regression of the set produced for this chapter. This model was produced and displayed in the same way as model 2, (as the previous regression was used against regression 1) inclusive of all vegetation data. The results from these models are shown in Table 4.8. The results from this last model are indeed surprising. All weed standardised coefficient regressions (when including all respiratory medications) are below -0.1, apart from mugwort (0.01), and dock, which instead has a positive value above 0.1. The regression coefficients for most of the different tree species indicate an inverse association with prescribing rates administered, including birch, although it should be of note that other species display larger coefficient values. Alder is the only tree species indicative of a positive association, which has the potential of suggesting higher prescribing rates in areas with a higher alder density. This is a point best covered in detail in the discussion, however. These patterns are observed with both all respiratory medications and just antihistamines.

Some final results of note from this regression are that both population age groups demonstrated positive associations in all findings, but the urban-rural categories were somewhat more complex: when all medications are considered, both groups have positive values with every vegetation but 3 types of weeds (Dock/Plantain/Ragweed), but when looking at only those in antihistamines, only mugwort contains both values as positive standardised coefficients. All other weeds have both categories for this as negative, and the grass and tree vegetation have positive regression coefficients for the town/fringe category compared to urban, but then negative when comparing urban to rural, which will make an interesting discussion point in the discussion. Most of the p-values were below 0.001, however, some values

were above the 0.05 threshold: considering all medications, dock weed in town/fringe areas produces a p-value of 0.358, plantain 0.015, and ragweed a value of 0.36. Looking at the regressions for antihistamines, the prescription rate, when regressed against the vegetation data for grass produces a p-value of 0.015, and 0.442 when run alongside ash, dock weed produces a value of 0.025 when run against ozone, and once again involving the town/fringe category, mugwort has a p-value of 0.012. These values will be discussed further in chapter 5 in subsequent sections for the corresponding vegetation and confounder variables.

Table 4.9 on page 57 is the final table in the chapter, and summarises the main results collected on the analyses run within this chapter. Here the coefficients are shown as either positive, negative, or not significant (NS). In the case of coefficients collected for testing the vegetation cover both during and out of season, only grass, ragweed, and birch were examined in this way.

Before moving to the discussion section, the goodness of fit should be commented on within the models. Models 1 and 6, unadjusted for the many variables returns an  $r^2$  value of less than 0.01 for each regression, indicating that less than 1% of the variation observed in the prescribing rates can be explained through the vegetation cover. However, upon examining the regression models where potential confounding variables have been adjusted for (Models 2-5 and 7), this value rises significantly, with  $r^2$  values mostly between the range of 0.4-0.5. This would indicate that the full models are able to explain 40-50% of the prescribing rate variation, although the majority of this isn't associated with vegetation cover variation.

**Table 4.8 – Regression Model 7: Grass and weeds mutually adjusted and adjusted for potential confounders**

Respiratory Medications					
Predictor	Coefficient	Coefficient (Std.)	95% Co		P> t
Grass	-0.033	-0.012	-0.044	-0.022	<0.001
Dock	19.777	0.143	19.017	20.536	<0.001
Mugwort	-45.574	-0.015	-61.243	-29.904	<0.001
Nettle	-18.709	-0.123	-19.596	-17.822	<0.001
Plantain	-11.160	-0.149	-11.595	-10.725	<0.001
Ragweed	-374.121	-0.144	-389.422	-358.819	<0.001
Alder	0.032	0.025	0.028	0.037	<0.001
Ash	-0.006	-0.009	-0.008	-0.004	<0.001
Birch	0.002	0.006	0.001	0.004	<0.001
Hazel	-0.008	-0.026	-0.009	-0.007	<0.001
Oak	0.008	0.012	0.006	0.010	<0.001
Pine	0.013	0.046	0.012	0.014	<0.001
Willow	-0.013	-0.011	-0.017	-0.009	<0.001

Antihistamines					
Predictor	Coefficient	Coefficient (Std.)	95% Co		P> t
Grass	0.007	0.005	0.001	0.013	0.015
Dock	11.368	0.152	10.976	11.761	<0.001
Mugwort	-70.666	-0.042	-78.768	-62.565	<0.001
Nettle	-11.766	-0.142	-12.224	-11.308	<0.001
Plantain	-6.695	-0.165	-6.920	-6.470	<0.001
Ragweed	-225.916	-0.160	-233.819	-218.014	<0.001
Alder	0.016	0.022	0.013	0.018	<0.001
Ash	0.000	-0.001	-0.001	0.001	0.442
Birch	0.004	0.018	0.003	0.005	<0.001
Hazel	0.000	-0.002	-0.001	0.000	0.217
Oak	0.007	0.019	0.006	0.008	<0.001
Pine	0.009	0.060	0.009	0.010	<0.001
Willow	-0.009	-0.013	-0.011	-0.007	<0.001



### Respiratory Medications

	Vegetation	Coefficients			
		Unadjusted	Adjusted	During Season	Outside Season
Weeds	Grass	+	-	-	-
	Dock	+	+	/	/
	Mugwort	-	-	/	/
	Nettle	-	-	/	/
	Plantain	-	-	/	/
Trees	Ragweed	-	-	-	-
	Alder	+	+	/	/
	Ash	-	-	/	/
	Birch	-	+	+	+
	Hazel	-	-	/	/
	Oak	-	+	/	/
	Pine	NS	+	/	/
	Willow	-	-	/	/

### Antihistamine Medications

	Vegetation	Coefficients			
		Unadjusted	Adjusted	During Season	Outside Season
Weeds	Grass	+	NS	NS	+
	Dock	+	+	/	/
	Mugwort	-	-	/	/
	Nettle	-	-	/	/
	Plantain	-	-	/	/
Trees	Ragweed	-	-	-	-
	Alder	+	+	/	/
	Ash	-	NS	/	/
	Birch	-	+	+	+
	Hazel	-	NS	/	/
	Oak	-	+	/	/
	Pine	-	+	/	/
	Willow	-	-	/	/

## 5 Discussion

With the results now presented and described in the tables and figures in the previous chapter, the next step is to discuss these findings and analyse what they show. Many variables of different measurements exist within each model, and within each many relationships and interactions can be identified and interpreted. This chapter will investigate these variables and their outcomes, deciphering the findings, determining what they suggest of the data and hypotheses, and in turn allowing conclusions to be drawn. Associations between prescribing and the vegetation measures will be the forefront of discussion, with particular interest in grass, birch and ragweed. After this associations with the rest of the vegetation types will be scrutinised, comparing it to those of interest and the other measures. Air pollutant data will be another major discussion point, with particular interest in particulate matter. Finally, the strengths and limitations of the study will be described, investigating the processes undertaken and further improvements and work that could be undertaken as a result of this study. This will be concluded in the next and final chapter, which will encompass this information and act as a closing statement to the study.

### 5.1 Grass Vegetation

Grass vegetation is most probably the vegetation of highest importance in the study. Grass pollens are considered one of the most allergenic species, and well renowned for their effects on individuals with rhinitis, and known interactions with confounders such as air pollution (D'Amato, Cecchi et al. 2007). Even with the added species, the variables and values collected for grasses are certainly one of the most interesting open to discussion following the models and results shown. From the initial descriptive statistics, although little data from grass vegetation exists within this figure, from what is present, the vegetation shows the species ranges from being absent altogether in any LSOA to as high as 87%, as discussed within the previous chapter. It can be assumed that LSOAs containing no grass whatsoever are some of the most urban areas within the country, with the highest percentages being some of the most rural.

Moving on from this, the first two regression models provide more information relevant for analysis. The positive regression coefficients in model 1 (unadjusted) compared to the lesser and negative coefficients in model 2 (adjusted) shows that some of the potential confounders do in fact appear to confound the association between the prescription rates and grass density.

The first two models, showing the coefficients unadjusted and subsequently adjusted for, show positive standardised coefficients for nearly all the grass values. It should be

acknowledged that the values when adjusted for the confounding variables are much lower, and when including all respiratory medications, the coefficient for grass is in fact negative, whilst remaining positive for antihistamines. The confounding variables are again adjusted for when considering the times within and out of the pollen seasons. Here, the coefficients are again lower than model 1, but follow the trend observed in model 2; when all respiratory medications are taken into consideration, negative associations are seen at all times of the year, but a positive association with the prescription rate is displayed with antihistamine medications. This is an interesting set of results to consider, as it does show to some extent what the literature shows, particularly with the results collected for antihistamines, but not when all medications are included in the regression (Erbaş, Chang et al. 2007, McInnes, Hemming et al. 2017). There are a few suggestions as to what could be responsible for these results differing from previous studies, such as the fact that not all medications administered are for grass vegetation, and that other confounders are at work. This is discussed below, and again in more detail in the relevant subsections.

In particular the air pollutants measured show interesting results. The negative coefficients suggest the pollutants ozone, nitrogen dioxide, and particulate matter are associated with lower prescription rates for respiratory conditions, whilst sulphur dioxide shows a positive association. Interestingly these results presented in this way contradict a number of papers concerning air pollution and pollens, which instead find that increased air pollution has a detrimental effect on those with allergies and rhinitis to pollens, including grasses such as Poaceae (D'Amato 2011, Pasqualini, Tedeschini et al. 2011). However, the context of the results for air pollutants can be thrown into doubt, such as how the particulate matter variable was examined in regression model 5. This will be discussed in greater detail in the subsection regarding air pollution (section 5.4).

Looking at the other variables in model 2, for both respiratory and antihistamine medications, the rest of the coefficients at least fit with other research conducted. These coefficients are almost entirely positive (with the exception of rural areas for antihistamine medications), and highlight several particular variables that seemingly have a significant role to play: the strong positive values of deprivation indices suggest that populations in more deprived areas are subject to higher prescription rates, whilst the population results show a positive association with both population age groups – much more so with elderly populations than those under 14 – suggesting an association between the prescription rates and both groups.

When comparing the results for models 3 and 4 between each other, detailing the changes between pollen seasons and the rest of the year, the results show that considering all

respiratory medications, the changes between models are slight, though some of the confounding variables do show some significant drops in association out of the pollen season. This will be covered further in the relevant subsections. As for the prescribing rates themselves, again an inverse association is observed through respiratory medications, and positive with antihistamines, as discussed above. However, when comparing only the rates during the pollen season with those out of season, the differences are so close it is not possible to say with certainty if there are significant association differences. The confounding variable differences may be the result of this, but this again, will be discussed in the relevant subsections concerning them.

This first part of the discussion has shown some agreement, and some disagreement with other data and research conducted previously. This has interesting implications; however, it must be acknowledged, and will continue to be throughout this chapter, that the data and the methods of examination are subject to limitations, which are discussed below in 5.6.

## **5.2 Ragweed and Birch**

### **5.2.1 Ragweed**

Ragweed pollens are considered one of the most significant of recent years in terms of allergenicity. Found particularly in urban areas, or areas with higher levels of pollution, these weeds produce pollen during seasons peaking in the late summer months (Taramarcaz, Lambelet et al. 2005). Based on the descriptive statistics, ragweed cover is almost negligible in density when compared to other vegetation, such as grass, but the results detail high regression coefficients with some variables.

The statistics for ragweed showed some interesting topics to open up discussion: to start with, the coefficients for prescription rates in models 1 and 2 show a negative association, which compared to grass and to the predictions for the data, is not what was expected. The unstandardised coefficients are in fact some of the largest differences in this data. Why this is the case is not understood, but could be as a result of some of the confounding variables, as the literature has shown the weed species are still problematic for many sufferers of asthma and allergies. As mentioned in the results, ragweed vegetation vs prescription rates for both types of medication indicate inverse associations, which is surprising, and the p-values indicate no reason to reject the null hypothesis. This is unexpected, as ragweed is considered highly allergenic, and on the rise in many countries around the world (Taramarcaz, Lambelet et al. 2005, Hamaoui-Laguel, Vautard et al. 2015). Yet this suggests this is not the case here, or at least

not in patients attending their general practitioner for medications to treat their symptoms. This is interesting, given how this goes against a lot of what the literature suggests, but is not necessarily the most reliable way of accurately examining the incidences of those allergic to pollens/requiring treatment for their symptoms (see section 5.6).

### **5.2.2 Birch**

The results of the descriptive statistics show birch trees are found in plentiful numbers in some LSOAs. The maximum number found in some LSOAs is in fact the third highest of any of the trees, as well as the highest mean value of any tree species. The high numbers of birch trees found in some LSOAs within Table 4.1 suggests there is an affinity for large numbers of birch to be found in rural areas. However, there are issues with categorising a continuous variable such as this, and the logistics involved in classifying land also throws some doubt to the results – areas of forest and farmland may both fall under one category, for example, yet one will contain significantly more birch vegetation than the other.

The initial unadjusted results show that for both types of medication, there is a negative association between birch vegetation and prescription rates. However, after adjustment for potential confounders (model 2) these coefficients change sign, and indicate a positive association in prescribing rates in areas with a higher birch tree density. These results are consistent with what might be expected if birch pollen is associated with higher rates of allergy and asthma. It must be stated that this is exclusive of months in which the pollen season is active/inactive, which is depicted in models 3 and 4. In model 3, the coefficients suggest a positive association with prescription rates and birch, which is reinforced by the literature, and therefore encouraging (Asam, Hofer et al. 2015, Grundström, Dahl et al. 2017). However, the same results are observed in model 4, in which only the times out of the pollen season are recoded, and as such suggest other factors are at play. In fact, a lot of p-values in model 4 are much higher than the threshold for birch, though interestingly not for model 3. This means it would be difficult to compare the two, given the validity of the values for prescription rates and birch vegetation cover out of its pollen season can be questioned. Further discussion about urban/rural as a confounder variable/effect modifier will be carried out in subsection 5.5.

### **5.2.3 Comparison across vegetation types**

This subsection will focus on the standardised coefficient models, which when examined in each model allow comparisons of the magnitude of associations between the prescribing rates and the three vegetation types. Starting with respiratory medications, when the confounding variables have not been adjusted for, there are some noted comparisons between the vegetation and prescriptions: grass coefficients are approximately half of the magnitude of that of ragweed, whilst birch values are the same as grass, but positive rather than negative. With the variables adjusted for, models 2-4 show similar standardised coefficients with each of the vegetation. Comparing the species, the magnitude of grass associations is approximately one twelfth that of ragweed, whilst birch is approximately one third that of grass (though once adjusted for variables, the birch values are positive rather than negative). Moving to antihistamine medications, again starting without any adjustment, grass and birch both share similar values, again with grass positive and birch negative. With the confounders included in the model calculations, grass and birch both return positive values, with ragweed again negative. When observing these values, the associations with grass are approximately a third of that with birch. Ragweed, on the other hand, is roughly fifty times greater than for grass.

### **5.3 Other Vegetation**

In this section, associations between prescribing and the rest of the vegetation types are discussed. Exploratory analyses were undertaken, as opposed to the more detailed investigation of the three key vegetation types of interest, and so the discussion is briefer than that for grass, ragweed, and birch. The interest here will not necessarily be on the species themselves, but rather comparisons between species in the models, and the ones of specific interest chosen from each type (ragweed from weed and birch from tree). This will determine if any of the further species examined are of significance, and how results compare to the other two.

#### **5.3.1 Weeds**

The four other weed species examined in parallel with ragweed were only done so using simple regressions both adjusted and unadjusted for the confounders. From the initial descriptive statistics, the other weed species are far more abundant than ragweed, although even the weed with the highest maximum levels in an LSOA (plantain) only reaches 2%. It is also

interesting to see that ragweed is the only weed species that is entirely absent in some LSOAs – every other is present in a small percentage at least.

The results of model 6 detailing the regression coefficients of all plants unadjusted for any other variables are the next table to be included in discussing the weeds. This table shows some interesting results: the coefficients for dock are in fact the only values that suggest a positive relationship between the weeds and the prescription rates. Ragweed shows a negative relationship in its coefficient values for both respiratory medications and antihistamines, whereas dock weed is positive for both types of medications, and both unaltered and standardised coefficients. The other weed species, however, show only negative relationships with the prescription rates in unadjusted models. Mugwort weeds return some of the largest negative relationships. However, plantain also shows a strong negative coefficient (only in comparison to the other weed values) when respiratory medications are considered.

The literature has much to say about the various weed species under scrutiny here. Primary sources suggest that whilst dock weed is the only species to yield a positive relationship against the health outcomes estimated by the prescription rate in our data, there is little evidence elsewhere to suggest that there is much danger of rhinitis or asthma resulting from the pollens of dock weed (McInnes, Hemming et al. 2017). As for the other species we have processed in the data, interestingly, mugwort appears to be one of the most allergenic species of those examined in previous research. However, the results here do not reflect this. There are potential reasons why these weren't picked up, which will be discussed momentarily. As with mugwort, similar findings can be found in primary literature relating to nettle pollen, suggesting a link between nettle pollens and rhinitis and asthma (Jeong, Son et al. 2016, Tiotiu, Brazdova et al. 2016, McInnes, Hemming et al. 2017). Finally, plantain weeds once again follow this trend, showing the opposite of what these results should if they were to follow the trends seen in the literature (De Weger, Bergmann et al. 2013, Gadermaier, Hauser et al. 2014).

The final model (model 7) includes all vegetation types and adjusts for confounding variables. Examining the differences between the models shows the trend between vegetations: once adjusted for the confounding variables, dock values are more positive in model 7 when adjusted for the variables than unadjusted in model 6, whilst the mugwort coefficient is greatly reduced in its negativity. Plantain and ragweed also return negative coefficients. The only plant species that is different when considering antihistamines instead of all respiratory medications is nettle, where the association with the prescription rate remains negative, but the magnitude itself increases. As well as this, though the values for antihistamines have consistently been closer to zero than respiratory medications (an understandable comparison, given the

broadness of one group and the specific nature of the other), once the confounding variables have been accounted for, differences between the two medication groups are much less visible. These changes in values are indicators of the potentially confounding variables in this study showing some relationships with the prescription rates and the vegetation densities.

### 5.3.2 Trees

Again, more exploratory analyses were undertaken with the tree data. Data for all six tree species are presented in Table 4.1, and the results of models 6 and 7 will be covered here, and compared to what has already been shown with birch trees. Starting with the descriptive statistics, whilst birch has the highest mean density (trees per km<sup>2</sup>), ash, oak, and pine are the next highest values, with willow and then alder both significantly lower. All tree species have at least one LSOA where each are absent entirely. Alder and willow trees have the lowest maximum sum values, while hazel and pine are by far the largest values, with at least four and nine times higher than the values for alder and willow respectively. Birch has the third highest maximum value by comparison.

Model 6 shows some interesting associations between prescribing and tree densities. Compared to weeds, the results of this model are a lot simpler: Alder is the only tree species that shows a positive relationship with prescription rates. The rest are all negative, a trend that is consistent with both types of medications investigated throughout the models. All p-values fall below the conventional 0.05 threshold, allowing the null hypothesis to be rejected. Examining the negative values in more detail amongst the standardised coefficients, ash and hazel trees show the most negative coefficient values with little difference between the two. Oak and pine tree densities are the closest values to zero, while birch and willow tree coefficients sit between the two, also with very close values between them. Again, the associations are similar for both medication types. Although this does not include any adjustment for confounding variables, this data provides an interesting insight: suggesting that only alder trees indicate a positive relationship between prescription rates and the species' density, whilst the other species show the inverse effect. Whilst this is not what has been shown in literature concerning tree pollens and the health outcomes investigated in this study, it must be stressed that this is unadjusted for confounders, which may have an effect on the results here (De Weger, Bergmann et al. 2013, McInnes, Hemming et al. 2017).

As with the weeds in the previous section, there are some interesting points to be made of regression model 7, particularly when comparing it to model 6. After adjustment, the



associations remain the same with both medication groups. However, when examining tree species, birch, oak, and pine all join alder in presenting a positive value, whilst the remaining tree species remain negative. This is particularly important, given this suggests that with the confounders are adjusted for, there are positive associations between some species and prescription rates that are negative when examining without confounders in model 6. It is safe to say therefore, that after adjustment, associations between tree density and prescribing outline significant differences to the unadjusted model. Examining the values themselves, it would appear that the lowest coefficients, and ergo interactions with prescription rates and plant species are at their weakest when examining just antihistamines.

One further interesting point within model 7 is that when considering just antihistamines two of the trees – ash and hazel – show a coefficient value of 0. This could mean a great many number of things, given the number of confounders involved in the model, but these values are still increases on the values collected in model 6, implying that the confounding variables have an effect nevertheless. It must be reinforced that this in no way suggests there is no association whatsoever between prescription rates and these two tree species. In conclusion, there could be a number of factors at play responsible for the changes seen here, and as with other findings in this chapter, study limitations (5.6) may explain these unexpected results (De Weger, Bergmann et al. 2013, Asam, Hofer et al. 2015).

## **5.4 Effect Modification by Air Pollution**

As discussed both in the results chapter and in the above subsections, the values for the four air pollutants of interest do follow a consistent trend across the regression models. Although these are interesting, they are not directly relevant to the research question, and as such the coefficient values produced in models 2, 3, and 4 will be discussed alongside the other confounder variables in section 5.5. This section will focus solely on the results produced by regression model 5, which examines how different levels of particulate matter affect the relationship between the prescription rates of medications and the three main vegetation types of interest.

### **5.4.1 PM<sub>10</sub>**

Particulate matter, as discussed in the literature review, is considered one of the most harmful air pollutants, with high levels found particularly in urban environments. As with the

other air pollutants analysed here, its direct relationship with the disorders asthma and rhinitis is more damaging the airways leading to dispositions to onsets of respiratory conditions than interactions with pollens and subsequent allergic responses. With pollen, the interaction can act in several ways, including air pollution interacting with and weakening the pollen grains, releasing the allergens within should the grain rupture (particularly particulate matter (Sénéchal, Visez et al. 2015)), and physical damage to the plants themselves, resulting in changes affecting the release of pollens in these plants. Particulate matter was chosen to be analysed over the other three pollutants due to its high levels in urban environments, and the known effects of interactions with pollens, and individuals suffering respiratory conditions (D'Amato 2011, Gleason, Bielory et al. 2014). PM<sub>10</sub> was at the forefront of examination in the regression model 5, investigating any potential interactions between the air pollutant and the three pollens of particular interest.

Tests for interactions described in chapter 4 indicated evidence of effect modification by PM<sub>10</sub> with the three key vegetation types. These were investigated by stratified analyses presented in Table 4.6. By splitting the continuous PM<sub>10</sub> variable into discrete categories, it is possible to look for any variation in the relationship between the plant species and prescribing, across levels of particulate matter. The results showed some interesting findings: The overall association between grass and antihistamine prescribing gave an unstandardized coefficient of -0.033 ( $p < 0.001$ ). However, when stratified by PM<sub>10</sub> quintile, the models suggested positive associations at higher concentrations of PM<sub>10</sub>. With ragweed a negative relationship is shown for both prescription rates, and birch does not display any distinct pattern between the vegetation and prescribing across quintiles for either prescription rate. These suggest, particularly with grass, that the overall association masks a more interesting result, which is more consistent with the hypotheses and suggestions that the vegetation impacts are greater at higher PM<sub>10</sub> concentrations.

In summary, grass density indicates associations with prescribing in interactions with PM<sub>10</sub> that are consistent with the hypotheses, however, this is not the case with birch and ragweed. The literature surrounding PM<sub>10</sub> suggests that particulate matter only seems to become detrimental to individuals at high levels, and in some cases no association is seen at all, or only in some environments (Takizawa 2011, Grundström, Dahl et al. 2017). This is interesting, as our results show a decrease in correlation between the prescription rate and ragweed, and positive with grass. However, given the various methods of interaction possible between the pollutant and the individual, further study is recommended to uncover the exact implications of the findings here. However, the p-values of these models, particularly with antihistamine

medications when examining birch vegetation spread does invite doubt. The literature backs this, with suggestions that factors such as weather can have a large effect on the vegetations and their pollens. There are also some further issues with the data that may be accountable for these results differing from what else has been shown, which will be discussed in full in section 5.6.

## **5.5 Associations with confounders**

This next section will examine the potential confounding variables in the study, consisting of confounders to the variables relevant to the outcomes of the project. Whilst these are therefore not the main focus of the study, it is still important to observe any associations with the prescribing rates.

### **5.5.1 Ozone, Nitrogen Dioxide, and Sulphur Dioxide**

Although particulate matter was examined with particular scrutiny, the other three pollutants were only measured as potential confounding variables within the models. The most interesting thing to note of these results – as mentioned above with particulate matter – is that there was little change in values for the air pollutants between different species models. Further similarities between models are the trends – ozone and nitrogen dioxide are also significantly negative when associated with prescription rates and vegetation, whilst sulphur dioxide remains the only pollutant studied to return a positive association. Again, there are several reasons as to what could be responsible for this, which will be discussed alongside findings and comparisons for each of the three.

Starting with ozone, the negative results go against what is known about the association between the pollutant and the health conditions. Nitrogen dioxide also shows the same association as ozone, and is the pollutant with the strongest negative associations with prescribing, again contradicting current evidence. It is known as one of the most potent pollutants, but is only found in high quantities within urban and industrial environments, with very little presence further out into the fringe and rural areas (Ribeiro, Costa et al. 2017). The reasons for these contradictory findings is cannot be explained by these results alone, but again this is not a focus for this study.

The final pollutant to acknowledge in this chapter, sulphur dioxide, returns consistently high coefficients, suggestive of a strong positive relationship with the prescription rate – the

only pollutant to display such an association. Interestingly enough, sulphur dioxide, like nitrogen dioxide, can be found in higher concentrations in urban and industrial areas, yet displays the opposite findings of those seen with NO<sub>2</sub>. Though there is a possibility that given sulphur dioxide can induce (reversible) respiratory difficulties at high concentrations, this association could potentially be explained in part by those suffering from short term exposure to the pollutant, as once the pollutant source has been removed, those affected have been shown to swiftly return to good health (D'Amato 2011).

### **5.5.2 – Urban/Rural Context**

The variable that accounts for the urban/rural environment of each General Practice is measured in a different way to the other (linear) variables in this study. The coefficients for this categorical variable represent the difference in prescribing rates between rural or town and fringe GPs and the urban category. Positive coefficients suggest a higher prescription rate in the fringe or rural areas, with a negative coefficient instead implying GPs in urban areas have higher prescription rates. The urban/rural environment was measured in all models looking at the three vegetations of interest and accounting for confounding variables (models 2-4). The descriptive statistics show that general practices in urban environments compose the majority of those in England, with just over 10% of that number on the rural-urban fringe, and approximately 4% in rural areas.

The coefficients comparing prescribing rates in more rural and urban settings are highly variable; they are often quite large and vary between positive and negative values depending on the vegetation type in the model and the medication group. Why this could be opens up a number of possibilities, such as the location of vegetation within an environment, and also interactions between them and other confounders. By their nature, more rural areas will also tend to have higher vegetation densities, so the results may reflect this. There are also limitations of using a simple three-level variable to capture urban/rural differences.

### **5.5.3 – IMD**

The Index of Multiple Deprivation (IMD) is a measure to calculate how deprived an area is. Taking into account income, employment, health, crime, education, and the local environment. Through this, the areas considered the most deprived in England can be identified, as well as those least deprived. The lower values in this variable detail those areas where deprivation is less, whilst the higher values indicate more deprived areas. It is expected that the

results should show more deprived areas at higher risk of asthma and rhinitis and therefore with higher prescribing rates. Throughout the results, the IMD variable continuously returns a strong positive coefficient value. There is always a positive relation with more deprived areas and the rate of prescriptions given. This is consistent with what might be expected, given the existing evidence on socio-economic deprivation and health (Cesaroni, Farchi et al. 2003).

#### **5.5.4 – Demographics – older and younger populations**

The population age groups are also interesting to examine, given the incidence rates of asthma and rhinitis are highest in children, yet the elderly are also susceptible, given the weakened immune system at an older age. Model 2 shows that it is the elderly age group that results in the highest prescription rates, though the age group for under 15s is also positively associated. Models 3 and 4 mirror this comparative trend, though show no true difference in rates during or out of pollen seasons. It is expected that elderly individuals will visit their general practitioners rather than other means of medical assistance, and although the incidence rate of asthma in children is higher, varying responses and severities in reactions to vegetations and other stimuli could prompt more urgent medical help than seeing a practitioner for medication (Pinto and Jeswani 2010). If this is the case, then the findings are not unusual to observe.

### **5.6 – The Study**

This final section will bring all the topics of discussion previously covered in this chapter together to examine the study in its entirety. Here the hypotheses, other findings, strengths and weaknesses of the project will all be brought in and discussed. Finally, this section will also look ahead at the future capabilities of this project, the data, and the impacts this could have.

#### **5.6.1 – Summary of findings**

To summarise the results in general, it can be concluded that there is no clear answer to our research question. The findings of the different analyses present mixed results, which cannot accept or reject either hypothesis. Some aspects of the project do align with the  $H_1$  hypothesis. The examination of  $PM_{10}$  with grass, for example is interesting and consistent with hypotheses. However, given the mixed results for many of the other findings and models within the study, this cannot be applied across the board. Some findings even go as far as to suggest inverse findings compared to what we would expect; implying that lower prescribing rates are

actually found in areas of high allergenic vegetation density, ragweed being the prime example. Given that some of the findings fit what is in the literature whilst others show the opposite, the one thing that can be concluded is that further study should be undertaken to address these uncertainties.

### **5.6.2 Strengths and Limitations**

As mentioned in previous sections in this chapter, this study has a number of limitations in design and data, the main ones of which will be discussed here. Before examining the data, it is important to acknowledge that in terms of design, this study is limited through the study being ecological. As a result of this it therefore lacks the power to infer cause and effect, since the data are aggregated across time and space. Unfortunately, this also limits how well confounding can be adjusted for. Should further analysis ever be carried out, it would be interesting to consider the possibility of using advanced panel data analysis to examine the dataset. This would allow for the time series of prescribing within each GP practice to be evaluated in relation to the vegetation data and other predictor variables. Turning to the data, we will start with the information regarding prescriptions administered at General Practices in England between June 2010 and December 2012. This data allows a monthly insight into the medications administered throughout the population. Although locations could be pinpointed on the practices, and estimates for population vegetation exposure calculated using variable radius buffers, there are a number of issues with this data that must be addressed here: In accordance with keeping the identity of all patients confidential, no information regarding the addresses of each patient was given. This means that accurate and detailed exposure of patients to specific vegetation types could not be made and we assume that the registered population lives within the somewhat arbitrary buffer zone. Particularly in town/fringe environments, where some patients may live in more urban settings and others more rural, there is potential to see with more detail how exposure can truly affect those based on location with finer detail. Another issue with the data is that there is no way to determine how often one individual may visit their GP for medication. Repeat visits cannot be shown with our data. This also means there is no way to determine if any patients have any susceptibilities to the vegetation or other confounders covered in this study that may result in increases or decreases in rates. Importantly, this data also does not allow the identification of the purpose for the medication, that is, the drivers for allergies that are not related to pollen. This is one of the main issues with the use of the data in this way; someone with an allergy to a stimulus different to pollens, such as food, pets, travel, or any other stimulus, would also be administered medication that falls under the inclusion criteria

used in this study (British National Formulary chapters, see chapter 3). Expanding on this, when including respiratory symptoms, there were a number of medications in this category that were also recommended for administration to individuals suffering conditions unrelated to allergies or rhinitis. COPD, for example, which is completely unrelated to our topic. One other downside of the sole use of this data, is that although many visit their GP for medications, as mentioned in the population subsection, many will seek more urgent medical attention should their responses prove faster and/or more severe. By the same token, many who suffer seasonal rhinitis may only experience mild symptoms, and as such may not need prescription strength medication, with many individuals likely obtaining medication for mild symptoms over the counter. By not having this data available for the study, a large percentage of individuals have been excluded that may distinctly affect the outcome of the findings in this project.

Looking at the other aspects, there are also a few issues that should be addressed within other parts of the data. The vegetation data, as one of the major pieces of secondary information in this study, is the next piece to acknowledge. Using average vegetation densities are subject to error in estimating population pollen exposure, given how factors such as weather can affect when pollens are released by plants, and the subsequent environments they affect. Although having an accurate representation of vegetation within each LSOA could be considered the next best thing, given the distance pollen grains can travel, and the different environments they may travel through and other potential confounders, there is a potential to suggest there are in some cases too many factors that can affect the grains to consider vegetation maps a very accurate exposure estimate. Air pollution is also a further aspect that can be mentioned here. Individuals with asthma and rhinitis can be affected by air pollution both directly and indirectly, with vegetation and their pollens not necessarily an essential part of the equation. As a result of this, it could be difficult to determine how much of a confounder air pollutants could be within this study.

Staying on the topic of measurements within each generated radius, the independence of observations needs to be mentioned. As can be seen in the example shown in figure 3.1 in the methods chapter, many GP practice buffers overlap, with some even encapsulating the entire radius of the smaller catchments in more densely-populated areas. Observation dependence was not taken into account, and it is possible the observed GP-level datapoints may not be entirely independent. Because observation dependence wasn't assessed or accounted for in the analyses, this is a limitation of the analytical approach taken.

The risks of collinearity are also an important point to acknowledge. Correlation coefficients between each key pair of variables at the GP practice level are shown in table 5.1 below. Whilst not affecting the reliability of the results as a whole, the risks of collinearity can affect the model results through one predictor being correlated with another. Table 5.1 shows the value of 0.05 when comparing grass with the prescription rate: a weak positive. However, the highest correlation between predictor values within the table is 0.857, for the predictor variables PM<sub>10</sub> and NO<sub>2</sub>. Although this is expected due to the common sources of these air pollutants (primarily fossil fuel combustion), future investigations could address this potential limitation of the analytical approach.

**Table 5.1 – Table of Correlation Coefficients between pairs of GP-level variables**

Colinear Annual Weighted Means (AWM)									
	Presc. Rate	Grass	Ozone	NO <sub>2</sub>	SO <sub>2</sub>	PM <sub>10</sub>	IMD	Population 0-14	Population 65+
Presc. Rate	1.000								
Grass	0.053	1.000							
Ozone	-0.150	0.236	1.000						
NO <sub>2</sub>	-0.204	-0.570	-0.556	1.000					
SO <sub>2</sub>	0.268	-0.240	-0.466	0.336	1.000				
PM <sub>10</sub>	-0.317	-0.536	-0.124	0.857	0.158	1.000			
IMD	0.356	-0.400	-0.528	0.469	0.375	0.253	1.000		
Population 0-14	0.100	-0.097	-0.089	0.159	0.146	0.160	0.342	1.000	
Population 65+	0.191	0.368	0.334	-0.600	-0.193	-0.521	-0.446	-0.463	1.000

The final limitation, as discussed in chapter 2, is that the urban/rural categorisation for areas is quite broad. A land use variable with capabilities of displaying more types of land/environments would allow for a much more thorough investigation, such a variable could provide insight into why birch does not change between pollen seasons. It is also imperative to acknowledge and understand the future impact of climate change and what this could mean, both for the health outcomes of those relevant to this study, but also for the changes in the pollen seasons brought on by these alterations, as well as changes to confounders such as air pollution, and the environments of the individuals, given the current and imminent threats invoked by this phenomenon (Lake, Jones et al. 2017).

These study limitations are balanced by its strengths. Using data collected from every General Practice in England also allows the study to be inclusive of all areas of the country and all individuals registered. Many previous studies have focused primarily on single places, and so



to encapsulate a picture of the full extent of a country is certainly a feat that has been proven successful in this project, and one that few previous studies have achieved. Another strength to draw upon involves the investigation of the role of confounders and effect modifiers, and studying the interactions and relationships of these that may affect the outcome variable. By taking into account and including these other variables, we can adjust for their role in the associations observed between the vegetation and prescription rates. Such findings have potential, for application in further study and for informing those attempting to manage their conditions, for whom such factors play a key role. The methods used to undertake the study have also proven a benefit: spatial softwares like Arc GIS and GME allow the data to be located with accuracy, and visually observed with ease. The synthesis of such variables as area weighted sums and means also allow a reasonably reliable estimate to be made of the vegetation spread within each LSOA. The vegetation maps should also receive a mention here. Data estimated for multiple vegetation types for every LSOA, most with a known allergenicity of varying magnitude, allows a thorough investigation across England, and for measurements to be undertaken within this study as to the exposure individuals within each LSOA are subject to. Another benefit of the study is the construction of the variable radius buffer used to estimate the vegetation, as well as the catchment area for each general practice. This allowed a good measurement of the area in which the population lived, and also the vegetation levels around each one, taking into account variable population diversity, as well as cross-LSOA values for vegetations, allowing comparisons across the country. Finally, the simple fact that many trends fit with not just the research question and  $H_1$  hypothesis, but with a lot of similar findings in previous studies shows promise in this project.

### **5.6.2 – Looking ahead**

Looking ahead, although a lot of findings have been produced by this work, and much can be drawn from these, there is still a vast amount of untapped potential from this study that could, and may be acted upon in the future. Ultimately the study did not clearly answer the core research question, and although some findings at least fell into what was expected in the project, the ambiguity of many of the results and the quantity of data means this study has the potential to inform future studies

Though the design of this study has allowed a lot of investigation and returned a great number of findings, there are many other study designs that if used, could further the work carried out in this study, as well as answering the research question. A key limitation of this study is that it is ecological. Using longitudinal cohort data could result in more robust results

for the study as a whole. Moving onto the prescribing data, it would be particularly interesting to determine how these compare with over the counter medication sales and admissions to Emergency Departments in hospitals over the same time periods. Such an examination of the data in the project would allow a more thorough investigation of the affected populations with more accuracy, and may even show findings as notably different to what was found in the project. It would also be fascinating to obtain data on land use in more detail. As the whole country is investigated in this study, finding out the use of the land, be it farming, industry, commercial, residential, etc. would again improve the design, and could be used to map out predicted growth of vegetation, as well as involving variables such as air pollution. The final main point to make in improving the study would be to include all data at a monthly rate (e.g. replacing annual air pollution estimates with monthly measures). This would ensure any seasonal changes in data other than prescriptions was collected, and would give more accurate results for varying months, the vegetation, and their subsequent pollen seasons.

Other aspects of further study include such suggestions as examining the potential of these allergens in outdoor environments compared to indoor. Much has been investigated regarding allergenicity of stimuli between indoor/outdoor settings, as well as with respiratory conditions. Examining any change between the allergens and environments as well as the varying conditions of said environments could prove an interesting route of investigation. Another possibility for further work would be to examine patient's IgE receptor levels against the pollens and other confounders in their local environments. Examining the severity of the reactions could provide some interesting findings. It would also allow further examination into the population variables, possibly allowing an insight into why the elderly incidence is so high in our study when it is acknowledged unanimously that young children are the most affected.

## 6 Conclusion

This study has indicated that vegetation densities are associated with prescription rates for the medications administered by general practitioners for allergic rhinitis and asthma in England, although not always in the expected direction. The models show that there are associations with each of the key vegetation types (grass, birch, ragweed) in both positive and negative manners, for the coefficients of both medication groups investigated (antihistamines and a wider group of treatments for 'respiratory conditions'). Of the three key vegetation types scrutinised in this study, all were associated with prescription rates of these conditions, but not necessarily all in the same way, with strong, weak, positive and negative coefficients collected across the regression models. When comparing pollen seasons and non-pollen seasons for the three species, no substantial differences in associations was shown, suggesting that other unmeasured factors may be involved.

The confounders and effect modifiers too have shown to have associations with the prescribing rates. Although some are more distinct than others, it can be concluded that air pollutants, urbanisation, deprivation, and population age profile all play a role in the likelihood of populations receiving prescriptions for a health outcome fitting with the conditions investigated in this study. Air pollution, investigated as one of the most potent confounders and effect modifiers, particularly particulate matter, provided some insightful yet contradictory results, with many negative associations, contrary to expectations.  $PM_{10}$  interactions in model 5 however fitted with what was expected, though only for grass, suggesting more positive associations between prescribing and grass density at higher  $PM_{10}$  concentrations. The indices of multiple deprivation showed perhaps the strongest positive correlation with the prescribing rate, suggesting it is perhaps the most significant factor involved in determining which populations are the most vulnerable to health outcomes such as rhinitis and asthma. The population age profiles showed positive trends for both older and younger age groups suggesting that whilst the elderly appear more susceptible, their age group is not exclusive in this, with young children also affected more so than other populations. Finally, conclusions drawn from the urban/rural context show variable associations, but do appear to suggest that living in town/fringe areas may be more beneficial than urban environments. Certainly, this is what fits with the literature and what is known. However, which urban/rural context is most beneficial remains dependent on other confounders, in particular the vegetations in question.

This study has allowed the observation of various vegetation, medication, and confounder variations across the length and breadth of England, investigating associations to

determine relationships and interactions between them. In doing so we have examined every LSOA and General Practice, as well as every medication given out relevant to the project. Such a thorough investigation has produced some valuable findings, although subject to limitations of design and data. Overall, this study has produced mixed findings, some of which fit with the literature, and raises a number of questions and possibilities for further study and development, which if pursued, would help further research and understanding within the field. This could aid a great many individuals within the population who have health difficulties resulting from respiratory disorders such as allergies, rhinitis, and asthma.

## Acknowledgements

Thank you to Dr Rachel McInnes at the Met Office for the provision of the vegetation density estimate data that were used for the study. I also acknowledge the provision of data from the various secondary sources used. Analyses and interpretations are the responsibility of the author and not the data providers.

Special thanks to my Supervisors Dr Benedict Wheeler and Dr Richard Sharpe, and my mentor Dr Karyn Morrissey, for helping me through thick and thin and never giving up on me, particularly when I found myself way out of my depth, and always being on hand to see me through.

Finally a big thank you to everyone at the European Centre for the Environment and Human Health, for giving me the chance of a lifetime, and to have the time of my life while working on his degree, and making me feel like part of a family. I will miss my time there dearly and hope to one day be able to come back.

## Reference List

Allergyuk.org. (2016). "Allergy Statistics." Retrieved 5th April 2016, from <https://www.allergyuk.org/allergy-statistics/allergy-statistics>

Andrusaityte, S., R. Grazuleviciene, J. Kudzyte, A. Bernotiene, A. Dedele and M. J. Nieuwenhuijsen (2016). "Associations between neighbourhood greenness and asthma in preschool children in Kaunas, Lithuania: a case–control study." *BMJ Open* **6**(4).

Aronson, J. K. (2016). Poaceae A2. *Meyler's Side Effects of Drugs (Sixteenth Edition)*. Oxford, Elsevier: 841-843.

Asam, C., H. Hofer, M. Wolf, L. Aglas and M. Wallner (2015). "Tree pollen allergens—an update from a molecular perspective." *Allergy* **70**(10): 1201-1211.

Barnes, C. S., N. E. Alexis, J. A. Bernstein, J. R. Cohn, J. G. Demain, E. Horner, E. Levetin, A. Nel and W. Phipatanakul (2013). "Climate change and our environment: the effect on respiratory and allergic disease." *The Journal of Allergy and Clinical Immunology: In Practice* **1**(2): 137-141.

Barrett, S. C. (2002). "The evolution of plant sexual diversity." *Nat Rev Genet* **3**(4): 274-284.

Bates, D. V., M. Baker-Anderson and R. Sizto (1990). "Asthma attack periodicity: A study of hospital emergency visits in Vancouver." *Environmental Research* **51**(1): 51-70.

Bhugwant, C., M. Bessafi, O. Favez, L. Chiappini, B. Sieja and E. Leoz-Garziandia (2013). "High Contribution of Sea Salt Aerosols on Atmospheric Particles Measured at an Urban Tropical Location in Reunion Island." *Journal of Environmental Protection* **04**(08): 828-842.

BNF (2011). *British National Formulary* **61**. London, BMJ Group.

Bousquet, J., N. Khaltaev, A. A. Cruz, J. Denburg, W. J. Fokkens, A. Togias, T. Zuberbier, C. E. Baena-Cagnani, G. W. Canonica, C. Van Weel, I. Agache, N. Aït-Khaled, C. Bachert, M. S. Blaiss, S. Bonini, L. P. Boulet, P. J. Bousquet, P. Camargos, K. H. Carlsen, Y. Chen, A. Custovic, R. Dahl, P. Demoly, H. Douagui, S. R. Durham, R. G. Van Wijk, O. Kalayci, M. A. Kaliner, Y. Y. Kim, M. L. Kowalski, P. Kuna, L. T. T. Le, C. Lemiere, J. Li, R. F. Lockey, S. Mavale-Manuel, E. O. Meltzer, Y.

Mohammad, J. Mullo, R. Naclerio, R. E. O’Hehir, K. Ohta, S. Ouedraogo, S. Palkonen, N. Papadopoulos, G. Passalacqua, R. Pawankar, T. A. Popov, K. F. Rabe, J. Rosado-Pinto, G. K. Scadding, F. E. R. Simons, E. Toskala, E. Valovirta, P. Van Cauwenberge, D. Y. Wang, M. Wickman, B. P. Yawn, A. Yorgancioglu, O. M. Yusuf, H. Zar, I. Annesi-Maesano, E. D. Bateman, A. B. Kheder, D. A. Boakye, J. Bouchard, P. Burney, W. W. Busse, M. Chan-Yeung, N. H. Chavannes, A. Chuchalin, W. K. Dolen, R. Emuzyte, L. Grouse, M. Humbert, C. Jackson, S. L. Johnston, P. K. Keith, J. P. Kemp, J. M. Klossek, D. Larenas-Linnemann, B. Lipworth, J. L. Malo, G. D. Marshall, C. Naspitz, K. Nekam, B. Niggemann, E. Nizankowska-Mogilnicka, Y. Okamoto, M. P. Orru, P. Potter, D. Price, S. W. Stoloff, O. Vandenplas, G. Viegi and D. Williams (2008). "Allergic Rhinitis and its Impact on Asthma (ARIA) 2008\*." Allergy **63**: 8-160.

Carinanos, P. and M. Casares-Porcel (2011). "Urban green zones and related pollen allergy: A review. Some guidelines for designing spaces with low allergy impact." Landscape and Urban Planning **101**(3): 205-214.

Cariñanos, P. and M. Casares-Porcel (2011). "Urban green zones and related pollen allergy: A review. Some guidelines for designing spaces with low allergy impact." Landscape and Urban Planning **101**(3): 205-214.

Cesaroni, G., S. Farchi, M. Davoli, F. Forastiere and C. A. Perucci (2003). "Individual and area-based indicators of socioeconomic status and childhood asthma." European Respiratory Journal **22**(4): 619-624.

Christenhusz, M. J. M. and J. W. Byng (2016). "The number of known plants species in the world and its annual increase." 2016 **261**(3): 17.

Cohen, A. J., H. Ross Anderson, B. Ostro, K. D. Pandey, M. Krzyzanowski, N. Künzli, K. Gutschmidt, A. Pope, I. Romieu, J. M. Samet and K. Smith (2005). "The Global Burden of Disease Due to Outdoor Air Pollution." Journal of Toxicology and Environmental Health, Part A **68**(13-14): 1301-1307.

Costabile, F. and I. Allegrini (2007). "Measurements and analyses of nitrogen oxides and ozone in the yard and on the roof of a street-canyon in Suzhou." Atmospheric Environment **41**(31): 6637-6647.

D'Amato, G. (2011). "Effects of climatic changes and urban air pollution on the rising trends of respiratory allergy and asthma." Multidisciplinary Respiratory Medicine **6**(1): 28-37.

D'Amato, G., K. C. Bergmann, L. Cecchi, I. Annesi-Maesano, A. Sanduzzi, G. Liccardi, C. Vitale, A. Stanzola and M. D'Amato (2014). "Climate change and air pollution - Effects on pollen allergy and other allergic respiratory diseases." Allergo Journal **23**(1): 32-38.

D'Amato, G., K. C. Bergmann, L. Cecchi, I. Annesi-Maesano, A. Sanduzzi, G. Liccardi, C. Vitale, A. Stanzola and M. D'Amato (2014). "Climate change and air pollution: Effects on pollen allergy and other allergic respiratory diseases." Allergo J Int **23**(1): 17-23.

D'Amato, G., L. Cecchi, S. Bonini, C. Nunes, I. Annesi-Maesano, H. Behrendt, G. Liccardi, T. Popov and P. Van Cauwenberge (2007). "Allergenic pollen and pollen allergy in Europe." Allergy: European Journal of Allergy and Clinical Immunology **62**(9): 976-990.

D'Amato, G., L. Cecchi, M. D'Amato and G. Liccardi (2010). "Urban air pollution and climate change as environmental risk factors of respiratory allergy: An update." Journal of Investigational Allergology and Clinical Immunology **20**(2): 95-102.

D'Amato, G., S. T. Holgate, R. Pawankar, D. K. Ledford, L. Cecchi, M. Al-Ahmad, F. Al-Enezi, S. Al-Muhsen, I. Ansotegui, C. E. Baena-Cagnani, D. J. Baker, H. Bayram, K. C. Bergmann, L. P. Boulet, J. T. M. Buters, M. D'Amato, S. Dorsano, J. Douwes, S. E. Finlay, D. Garrasi, M. Gómez, T. Haahtela, R. Halwani, Y. Hassani, B. Mahboub, G. Marks, P. Michelozzi, M. Montagni, C. Nunes, J. J. W. Oh, T. A. Popov, J. Portnoy, E. Ridolo, N. Rosário, M. Rottem, M. Sánchez-Borges, E. Sibanda, J. J. Sienra-Monge, C. Vitale and I. Annesi-Maesano (2015). "Meteorological conditions, climate change, new emerging factors, and asthma and related allergic disorders. A statement of the World Allergy Organization." World Allergy Organization Journal **8**(1).

D'Amato, G., G. Liccardi, M. D'Amato and S. Holgate (2005). "Environmental risk factors and allergic bronchial asthma." Clinical and Experimental Allergy **35**(9): 1113-1124.

D'Amato, G., G. Liccardi and G. Frenguelli (2007). "Thunderstorm-asthma and pollen allergy." Allergy: European Journal of Allergy and Clinical Immunology **62**(1): 11-16.



Dadvand, P., C. M. Villanueva, L. Font-Ribera, D. Martinez, X. Basagaña, J. Belmonte, M. Vrijheid, R. Gražulevičienė, M. Kogevinas and M. J. Nieuwenhuijsen (2014). "Risks and Benefits of Green Spaces for Children: A Cross-Sectional Study of Associations with Sedentary Behavior, Obesity, Asthma, and Allergy." Environmental Health Perspectives **122**(12): 1329-1335.

Dales, R. E., S. Cakmak, S. Judek and F. Coates (2008). "Tree pollen and hospitalization for asthma in urban Canada." International Archives of Allergy and Immunology **146**(3): 241-247.

Davidson, A. C., J. Emberlin, A. D. Cook and K. M. Venables (1996). "A major outbreak of asthma associated with a thunderstorm: experience of accident and emergency departments and patients' characteristics. Thames Regions Accident and Emergency Trainees Association." Bmj **312**(7031): 601-604.

De Weger, L. A., K. C. Bergmann, A. Rantio-Lehtimäki, A. Dahl, J. Buters, C. Déchamp, J. Belmonte, M. Thibaudon, L. Cecchi, J. P. Besancenot, C. Galán and Y. Waisel (2013). Impact of pollen. Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts. **9789400748811**: 161-215.

Elholm, G., A. Linneberg, L. L. N. Husemoen, O. Omland, P. M. Gronager, T. Sigsgaard and V. Schlunssen (2016). "The Danish urban-rural gradient of allergic sensitization and disease in adults." Clinical and Experimental Allergy **46**(1): 103-111.

Erbas, B., J. H. Chang, S. Dharmage, E. K. Ong, R. Hyndman, E. Newbiggin and M. Abramson (2007). "Do levels of airborne grass pollen influence asthma hospital admissions?" Clin Exp Allergy **37**(11): 1641-1647.

Feo Brito, F., P. Mur Gimeno, C. Martinez, A. Tobias, L. Suarez, F. Guerra, J. M. Borja and A. M. Alonso (2007). "Air pollution and seasonal asthma during the pollen season. A cohort study in Puertollano and Ciudad Real (Spain)." Allergy **62**(10): 1152-1157.

Friedman, J. and S. C. H. Barrett (2009). "Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants." Annals of Botany **103**(9): 1515-1527.

Gadermaier, G., M. Hauser and F. Ferreira (2014). "Allergens of weed pollen: An overview on recombinant and natural molecules." Methods **66**(1): 55-66.

García-Mozo, H., C. Galán, J. Belmonte, D. Bermejo, P. Candau, C. Díaz de la Guardia, B. Elvira, M. Gutiérrez, V. Jato, I. Silva, M. M. Trigo, R. Valencia and I. Chuine (2009). "Predicting the start and peak dates of the Poaceae pollen season in Spain using process-based models." Agricultural and Forest Meteorology **149**(2): 256-262.

Gennaro, D. A., S. Antonello, L. Gennaro and D. A. Maria (2011). "Climatic Change and Urban air pollution on the rising trends of Respiratory Allergy and Asthma." Allergy, Asthma and Immunology: From Genes to Clinical Application: Proceedings of the Iv World Asthma and Copd Forum and Xvi International Congress on Rehabilitation in Medicine and Immunorehabilitation: 15-31.

Gilbert, N. (2016). "Green space: A natural high." Nature **531**(7594): S56-S57.

Girgis, S. T., G. B. Marks, S. H. Downs, A. Kolbe, G. N. Car and R. Paton (2000). "Thunderstorm-associated asthma in an inland town in south-eastern Australia. Who is at risk?" European Respiratory Journal **16**(1): 3-8.

Gleason, J. A., L. Bielory and J. A. Fagliano (2014). "Associations between ozone, PM2.5, and four pollen types on emergency department pediatric asthma events during the warm season in New Jersey: a case-crossover study." Environ Res **132**: 421-429.

Gonzalez-Barcala, F. J., J. Aboal-Vinas, M. J. Aira, C. Regueira-Mendez, L. Valdes-Cuadrado, J. Carreira, M. T. Garcia-Sanz and B. Takkouche (2013). "Influence of pollen level on hospitalizations for asthma." Arch Environ Occup Health **68**(2): 66-71.

Gonzalez Barcala, F. J., J. Aboal Viñas, L. Valdes Cuadrado, A. Bourdin, J. M. Alvarez Dobaño and B. Takkouche (2010). "Trends in hospital admissions due to asthma in north-west Spain from 1995 to 2007." Allergologia et Immunopathologia **38**(5): 254-258.

Grundström, M., Å. Dahl, T. Ou, D. Chen and H. Pleijel (2017). "The relationship between birch pollen, air pollution and weather types and their effect on antihistamine purchase in two Swedish cities." Aerobiologia.

Grunig, G., L. M. Marsh, N. Esmaeil, K. Jackson, T. Gordon, J. Reibman, G. Kwapiszewska and S.-

H. Park (2014). "Perspective: ambient air pollution: inflammatory response and effects on the lung's vasculature." Pulmonary Circulation **4**(1): 25-35.

Hamaoui-Laguel, L., R. Vautard, L. Liu, F. Solmon, N. Viovy, D. Khvorostyanov, F. Essl, I. Chuine, A. Colette, M. A. Semenov, A. Schaffhauser, J. Storkey, M. Thibaudon and M. M. Epstein (2015). "Effects of climate change and seed dispersal on airborne ragweed pollen loads in Europe." Nature Clim. Change **5**(8): 766-771.

Hess, J., L. Darrow, M. Klein, P. Tolbert and S. Sarnat (2010). "15: Impact of Daily Pollen Levels on Emergency Department Visits for Asthma." Annals of Emergency Medicine **56**(3, Supplement): S5-S6.

Hong, J., T. Zhong, H. Li, J. Xu, X. Ye, Z. Mu, Y. Lu, A. Mashaghi, Y. Zhou, M. Tan, Q. Li, X. Sun, Z. Liu and J. Xu (2016). "Ambient air pollution, weather changes, and outpatient visits for allergic conjunctivitis: A retrospective registry study." Scientific Reports **6**: 23858.

Hunt, J. and D. Eisenberg (2010). "Mental Health Problems and Help-Seeking Behavior Among College Students." Journal of Adolescent Health **46**(1): 3-10.

IPCC (2014). "IPCC AR5 Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change." IPCC Geneva, Switzerland: 151 pp.

Janke, K. (2014). "Air pollution, avoidance behaviour and children's respiratory health: Evidence from England." Journal of Health Economics **38**: 23-42.

Jeong, K. Y., M. Son, S.-Y. Choi, K. H. Park, H. J. Park, C.-S. Hong, J.-H. Lee and J.-W. Park (2016). "Standardization of Weed Pollen Extracts, Japanese Hop and Mugwort, in Korea." Yonsei Med J **57**(2): 399-406.

Johnston, N. W. and M. R. Sears (2006). "Asthma exacerbations · 1: Epidemiology." Thorax **61**(8): 722-728.

Kanter, U., W. Heller, J. Durner, J. B. Winkler, M. Engel, H. Behrendt, A. Holzinger, P. Braun, M. Hauser, F. Ferreira, K. Mayer, M. Pfeifer and D. Ernst (2013). "Molecular and Immunological Characterization of Ragweed (*Ambrosia artemisiifolia* L.) Pollen after Exposure of the Plants to Elevated Ozone over a Whole Growing Season." Plos One **8**(4).

Kasprzyk, I. (2008). "Non-native *Ambrosia* pollen in the atmosphere of Rzeszow (SE Poland); evaluation of the effect of weather conditions on daily concentrations and starting dates of the pollen season." Int J Biometeorol **52**(5): 341-351.

Kerr, R. A. (2007). "Global Warming Is Changing the World." Science **316**(5822): 188-190.

Kim, K.-H., E. Kabir and S. Kabir (2015). "A review on the human health impact of airborne particulate matter." Environment International **74**: 136-143.

Kraaijeveld, K., L. A. de Weger, M. Ventayol Garcia, H. Buermans, J. Frank, P. S. Hiemstra and J. T. den Dunnen (2015). "Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing." Mol Ecol Resour **15**(1): 8-16.

Kirkwood BR, Sterne JAC, Kirkwood BR. Essential medical statistics. 2nd ed. Oxford: Blackwell Science 2003.

Lake, I. R., N. R. Jones, M. Agnew, C. M. Goodess, F. Giorgi, L. Hamaoui-Laguel, M. A. Semenov, F. Solomon, J. Storkey, R. Vautard and M. M. Epstein (2017). "Climate Change and Future Pollen Allergy in Europe." Environmental Health Perspectives **125**(3): 385-391.

Lee, S. L., W. H. S. Wong and Y. L. Lau (2006). "Association between air pollution and asthma admission among children in Hong Kong." Clinical and Experimental Allergy **36**(9): 1138-1146.

Lu, S., J. Ren, X. Hao, D. Liu, R. Zhang, M. Wu, F. Yi, J. Lin, Y. Shinich and Q. Wang (2014). "Characterization of protein expression of *Platanus* pollen following exposure to gaseous pollutants and vehicle exhaust particles." Aerobiologia **30**(3): 281-291.

Makra, L., I. Matyasovszky, B. Balint and Z. Csepe (2014). "Association of allergic rhinitis or asthma with pollen and chemical pollutants in Szeged, Hungary, 1999-2007." Int J Biometeorol **58**(5): 753-768.

Marcazzan, G. M., S. Vaccaro, G. Valli and R. Vecchi (2001). "Characterisation of PM10 and PM2.5 particulate matter in the ambient air of Milan (Italy)." Atmospheric Environment **35**(27): 4639-4650.

Marieb, E. N. and K. Hoehn (2013). Human Anatomy & Physiology, Pearson.

Masoli, M., D. Fabian, S. Holt, R. Beasley and P. Global Initiative for Asthma (2004). "The global burden of asthma: executive summary of the GINA Dissemination Committee Report." Allergy **59**(5): 469-478.

Mathur, S. K. (2010). "Allergy and Asthma in the Elderly." Semin Respir Crit Care Med **31**(5): 587-595.

Matsuoka, R. H. and R. Kaplan (2008). "People needs in the urban landscape: Analysis of Landscape And Urban Planning contributions." Landscape and Urban Planning **84**(1): 7-19.

McInnes, R. N., D. Hemming, P. Burgess, D. Lyndsay, N. J. Osborne, C. A. Skjøth, S. Thomas and S. Vardoulakis (2017). "Mapping allergenic pollen vegetation in UK to study environmental exposure and human health." Science of The Total Environment **599**: 483-499.

Metoffice.gov.uk. (2016). "Pollen Count Calendar." Retrieved 18th August 2016, from <http://www.metoffice.gov.uk/public/weather/pollen-forecast/#?tab=map&map=Pollen&fcTime=1504004400&zoom=5&lon=-4.00&lat=55.71>.

Monks, P. S., C. Granier, S. Fuzzi, A. Stohl, M. L. Williams, H. Akimoto, M. Amann, A. Baklanov, U. Baltensperger, I. Bey, N. Blake, R. S. Blake, K. Carslaw, O. R. Cooper, F. Dentener, D. Fowler, E. Fragkou, G. J. Frost, S. Generoso, P. Ginoux, V. Grewe, A. Guenther, H. C. Hansson, S. Henne, J. Hjorth, A. Hofzumahaus, H. Huntrieser, I. S. A. Isaksen, M. E. Jenkin, J. Kaiser, M. Kanakidou, Z. Klimont, M. Kulmala, P. Laj, M. G. Lawrence, J. D. Lee, C. Liousse, M. Maione, G. McFiggans, A. Metzger, A. Mieville, N. Moussiopoulos, J. J. Orlando, C. D. O'Dowd, P. I. Palmer, D. D. Parrish, A. Petzold, U. Platt, U. Pöschl, A. S. H. Prévôt, C. E. Reeves, S. Reimann, Y. Rudich, K. Sellegri, R. Steinbrecher, D. Simpson, H. ten Brink, J. Theloke, G. R. van der Werf, R. Vautard, V. Vestreng, C. Vlachokostas and R. von Glasow (2009). "Atmospheric composition change – global and regional air quality." Atmospheric Environment **43**(33): 5268-5350.

Ogershok, P. R., D. J. Warner, M. B. Hogan and N. W. Wilson (2007). "Prevalence of pollen sensitization in younger children who have asthma." Allergy and Asthma Proceedings **28**(6): 654-658.

Okada, H., C. Kuhn, H. Feillet and J. F. Bach (2010). "The 'hygiene hypothesis' for autoimmune and allergic diseases: an update." Clin Exp Immunol **160**(1): 1-9.

Osthoff, H. D., J. M. Roberts, A. R. Ravishankara, E. J. Williams, B. M. Lerner, R. Sommariva, T. S. Bates, D. Coffman, P. K. Quinn, J. E. Dibb, H. Stark, J. B. Burkholder, R. K. Talukdar, J. Meagher, F. C. Fehsenfeld and S. S. Brown (2008). "High levels of nitryl chloride in the polluted subtropical marine boundary layer." Nature Geosci **1**(5): 324-328.

Packe, G. E. and J. Ayres (1985). "ASTHMA OUTBREAK DURING A THUNDERSTORM." The Lancet **326**(8448): 199-204.

Pasqualini, S., E. Tedeschini, G. Frenguelli, N. Wopfner, F. Ferreira, G. D'Amato and L. Ederli (2011). "Ozone affects pollen viability and NAD(P)H oxidase release from *Ambrosia artemisiifolia* pollen." Environmental Pollution **159**(10): 2823-2830.

Paul J. Brochu, S., Jeff D. Yanosky, ScD, Christopher J. Paciorek, PhD, Joel Schwartz, PhD, Jarvis T. Chen, PhD, Robert F. Herrick, ScD, and Helen H. Suh, ScD (2011). "Particulate Air Pollution and Socioeconomic Position in Rural and Urban Areas of the Northeastern United States." American Journal of Public Health **101**(S1): S224-S230.

Pawankar, R., G. Canonica, S. Holgate and R. Lockey (2011). "World Allergy Organization (WAO) white book on allergy." Wisconsin: World Allergy Organisation. Available online at: [http://www.worldallergy.org/UserFiles/file/WAO-White-Book-on-Allergy web. pdf](http://www.worldallergy.org/UserFiles/file/WAO-White-Book-on-Allergy%20web.pdf).

Peel, R. G., R. Kennedy, M. Smith and O. Hertel (2014). "Do urban canyons influence street level grass pollen concentrations?" International Journal of Biometeorology **58**(6): 1317-1325.

Petercel, R., L. Srnc, J. Culig, K. Zaninovic, B. Mitic and I. Vukusic (2004). "Atmospheric pollen season in Zagreb (Croatia) and its relationship with temperature and precipitation." Int J Biometeorol **48**(4): 186-191.

Pinto, J. M. and S. Jeswani (2010). "Rhinitis in the geriatric population." Allergy, Asthma, and Clinical Immunology : Official Journal of the Canadian Society of Allergy and Clinical Immunology **6**(1): 10-10.

Ribeiro, H., C. Costa, I. Abreu and J. C. G. Esteves da Silva (2017). "Effect of O<sub>3</sub> and NO<sub>2</sub> atmospheric pollutants on *Platanus x acerifolia* pollen: Immunochemical and spectroscopic analysis." Science of The Total Environment **599**: 291-297.

Ruffin, J., D. Williams, U. Banerjee and K. Pinnix (2009). "The effects of some environmental gaseous pollutants on pollen-wall proteins of certain airborne pollen grains: a preliminary study." Grana **22**(3): 171-175.

Sayin, İ. (2013). "An important social problem: allergic rhinitis." Journal of Medical Updates **3**(2): 91-95.

Schenker, M. (2005). "Farming and asthma." Occupational and Environmental Medicine **62**(4): 211-212.

Schüler, L. and H. Behling (2011). "Poaceae pollen grain size as a tool to distinguish past grasslands in South America: a new methodological approach." Vegetation History and Archaeobotany **20**(2): 83-96.

Scott, K. M., M. Von Korff, J. Ormel, M.-y. Zhang, R. Bruffaerts, J. Alonso, R. C. Kessler, H. Tachimori, E. Karam, D. Levinson, E. J. Bromet, J. Posada-Villa, I. Gasquet, M. C. Angermeyer, G. Borges, G. de Girolamo, A. Herman and J. M. Haro (2007). "Mental disorders among adults with asthma: results from the World Mental Health Survey." General Hospital Psychiatry **29**(2): 123-133.

Senechal, H., N. Visez, D. Charpin, Y. Shahali, G. Peltre, J. P. Biolley, F. Lhuissier, R. Couderc, O. Yamada, A. Malrat-Domenge, N. Pham-Thi, P. Poncet and J. P. Sutra (2015). "A Review of the Effects of Major Atmospheric Pollutants on Pollen Grains, Pollen Content, and Allergenicity." ScientificWorldJournal **2015**: 940243.

Sénéchal, H., N. Visez, D. Charpin, Y. Shahali, G. Peltre, J. P. Biolley, F. Lhuissier, R. Couderc, O. Yamada, A. Malrat-Domenge, N. Pham-Thi, P. Poncet and J. P. Sutra (2015). "A review of the effects of major atmospheric pollutants on pollen grains, pollen content, and allergenicity." Scientific World Journal **2015**.

Sheehan, W. J., P. A. Rangsinhienchai, S. N. Baxi, A. Gardynski, A. Bharmanee, E. Israel and W. Phipatanakul (2010). "Age-specific prevalence of outdoor and indoor aeroallergen sensitization in boston." Clinical Pediatrics **49**(6): 579-585.

Sih, T. and O. Mion (2010). "Allergic rhinitis in the child and associated comorbidities." Pediatr Allergy Immunol **21**(1 Pt 2): e107-113.

Smith, M., J. Emberlin and A. Kress (2005). "Examining high magnitude grass pollen episodes at Worcester, United Kingdom, using back-trajectory analysis." Aerobiologia **21**(2): 85-94.

Sousa, S. I. V., F. G. Martins, M. C. Pereira, M. C. M. Alvim-Ferraz, H. Ribeiro, M. Oliveira and I. Abreu (2008). "Influence of atmospheric ozone, PM10 and meteorological factors on the concentration of airborne pollen and fungal spores." Atmospheric Environment **42**(32): 7452-7464.

Stewart, I. D. and T. R. Oke (2012). "Local Climate Zones for Urban Temperature Studies." Bulletin of the American Meteorological Society **93**(12): 1879-1900.

Stone, K. D., C. Prussin and D. D. Metcalfe (2010). "IgE, mast cells, basophils, and eosinophils." Journal of Allergy and Clinical Immunology **125**(2, Supplement 2): S73-S80.

Takizawa, H. (2011). "Impact of Air Pollution on Allergic Diseases." The Korean Journal of Internal Medicine **26**(3): 262-273.

Taramarcas, P., C. Lambelet, B. Clot, C. Keimer and C. Hauser (2005). "Ragweed (Ambrosia) progression and its health risks: Will Switzerland resist this invasion?" Swiss Medical Weekly **135**(37-38): 538-548.



Tarlo, S. M., I. Broder, P. Corey, M. Chan-Yeung, A. Ferguson, A. Becker, C. Rogers, M. Okada and J. Manfreda (2001). "The role of symptomatic colds in asthma exacerbations: Influence of outdoor allergens and air pollutants." Journal of Allergy and Clinical Immunology **108**(1): 52-58.

Tiotiu, A., A. Brazdova, C. Longé, P. Gallet, M. Morisset, V. Leduc, C. Hilger, C. Broussard, R. Couderc, J.-P. Sutra, H. Sénéchal and P. Poncet (2016). "Urtica dioica pollen allergy." Annals of Allergy, Asthma & Immunology **117**(5): 527-534.

Tobías, A., I. Galán, J. R. Banegas and E. Aránguez (2003). "Short term effects of airborne pollen concentrations on asthma epidemic." Thorax **58**(8): 708-710.

Todea, D. A., I. Suatean, A. C. Coman and L. E. Rosca (2013). "The Effect of Climate Change and Air Pollution on Allergenic Potential of Pollens." Notulae Botanicae Horti Agrobotanici Cluj-Napoca **41**(2): 646-650.

Vardoulakis, S., E. Solazzo and J. Lumbreras (2011). "Intra-urban and street scale variability of BTEX, NO<sub>2</sub> and O<sub>3</sub> in Birmingham, UK: Implications for exposure assessment." Atmospheric Environment **45**(29): 5069-5078.

Venables, K. M., U. Allitt, C. G. Collier, J. Emberlin, J. B. Greig, P. J. Hardaker, J. H. Highham, T. Laing-Morton, R. L. Maynard, V. Murray, D. Strachan and R. D. Tee (1997). "Thunderstorm-related asthma--the epidemic of 24/25 June 1994." Clin Exp Allergy **27**(7): 725-736.

von Hertzen, L. and T. Haahtela (2005). "Signs of reversing trends in prevalence of asthma." Allergy **60**(3): 283-292.

Wardman, A. E., D. Stefani and J. C. MacDonald (2002). "Thunderstorm-associated asthma or shortness of breath epidemic: a Canadian case report." Can Respir J **9**(4): 267-270.

Wheeler, B. W., R. Lovell, S. L. Higgins, M. P. White, I. Alcock, N. J. Osborne, K. Husk, C. E. Sabel and M. H. Depledge (2015). "Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality." International Journal of Health Geographics **14**(1): 17.

Zhong, W., L. Levin, T. Reponen, G. K. Hershey, A. Adhikari, R. Shukla and G. LeMasters (2006). "Analysis of short-term influences of ambient aeroallergens on pediatric asthma hospital visits." Sci Total Environ **370**(2-3): 330-336.

Ziska, L. H., D. E. Gebhard, D. A. Frenz, S. Faulkner, B. D. Singer and J. G. Straka (2003). "Cities as harbingers of climate change: Common ragweed, urbanization, and public health." Journal of Allergy and Clinical Immunology **111**(2): 290-295.

## Appendix

### 1. Stata Do-file script

```
>*Do file to import and manage monthly GP prescribing data. Data source:
http://www.hscic.gov.uk/gpprescribingdata

>cd "E:\Projects\Pollen_LT\gp_prescribing\raw_HSCIC"

>*Loop to pull out monthly antidepressant prescribing data, and to append all months into a
single dataset

>set more off

*****

>foreach month in T201212 T201211 T201210 T201209 T201208 T201207 T201206 T201205
T201204 T201203 T201202 T201201 T201112 T201111 T201110 T201109 T201108 T201107
T201106 T201105 T201104 T201103 T201102 T201101 T201012 T201011 T201010 T201009
T201008 T201007 T201006

>*preserve the current data

>preserve

>*Clear current data from memory then import the prescribing data for the relevant month in
the loop,

>import delimited "`month'PDP IEXT.CSV", clear

>*Keep only antidepressant prescribing (from BNF chapter 3.1)

>keep if substr( bnfcode,1,4)=="0301" | substr( bnfcode,1,4)=="0302" | substr(
bnfcode,1,4)=="0303" | substr( bnfcode,1,4)=="0304" | substr( bnfcode,1,4)=="0305" | substr(
bnfcode,1,4)=="0308" | substr( bnfcode,1,6)=="030901"

>*Save a temp file, overwriting the previous month of data

>save temp ,replace

>*Restore the data from the preserve above

>restore

>*Append the next month's data using the temp dataset

>append using temp ,force
```

```

>*end the loop

>}

*****

>*Save the full appended dataset

>save prescribing_GP_England_allmonths_actual.dta

*Open the full appended dataset

>use prescribing_GP_England_allmonths_actual.dta, clear

*Preserve the data in memory

>preserve

* summed by GP practice and month = one observation per practice per month

>collapse (sum) items nic actcost, by(practice period)

*Save the full appended collapsed dataset

>save prescribing_GP_England_allmonths_all_selectedBNF_full.dta

>restore

*Keep only antihistamines/allergy drugs

>keep if substr( bnfcode,1,4)=="0301"

* summed by GP practice and month = one observation per practice per month

>collapse (sum) items nic actcost, by(practice period)

*Save the full appended collapsed dataset

>save prescribing_GP_England_allmonths_antihistamines_allergy_drugs.dta

>clear

*****

*Merge in populations and calculate prescribing rates. Full subset of prescriptions

>use prescribing_GP_England_allmonths_all_selectedBNF_full.dta, clear

```

\*Merge in 3 years of practice populations

>mmerge practice using

```
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2010_2011.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)
```

>mmerge practice using

```
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2011_2012.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)
```

>mmerge practice using

```
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2012_2013.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)
```

>gen prate\_all=1000\*items/Population\_10\_11 if period>201005 & period<201104

>replace prate\_all=1000\*items/Population\_11\_12 if period>201103 & period<201204

>replace prate\_all=1000\*items/Population\_12\_13 if period>201203 & period<201304

\*Drop unwanted variables

>drop nic actcost \_merge

\*rename items

>rename items items\_all

\*Save the file

>save prate\_all\_by\_practice\_period.dta

\*Antihistamines

>use prescribing\_GP\_England\_allmonths\_antihistamines\_allergy\_drugs.dta, clear

>\*Merge in 3 years of practice populations

>mmerge practice using

```
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2010_2011.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)
```

>mmerge practice using

```
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2011_2012.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)
```

```

>mmerge practice using
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_Pop_Age_2012_2013.dta" , type(n:1)
missing(nomatch) unmatched(master) umatch(Practice)

>gen prate_antih=1000*items/Population_10_11 if period>201005 & period<201104

>replace prate_antih=1000*items/Population_11_12 if period>201103 & period<201204

>replace prate_antih=1000*items/Population_12_13 if period>201203 & period<201304

*Drop unwanted variables

>drop nic actcost _merge

*rename items

>rename items items_antih

*Save the file

>save prate_antih_by_practice_period.dta, replace

*****

*Merge the two presc rate files together

>use prate_all_by_practice_period.dta, clear

>mmerge practice period using prate_antih_by_practice_period.dta, type(1:1)
missing(nomatch) unmatched(master)

>drop _merge

>la var prate_all "Prescriptions per 1000 pop, all subset drugs"

>la var prate_antih "Prescriptions per 1000 pop, antihistamines BNF3.1"

>save prate_all_antih_by_practice_period.dta, replace

*****

*Import GP addresses CSV file

>use prate_all_antih_by_practice_period.dta, clear

>import delimited
E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_England_allmonths_addresses_BNGxy.csv,
clear

```

```
>save imported_gp_addresses_dataset.dta, replace
```

```
*****
```

```
*Merge in GP addresses and OS co-ords
```

```
>use prate_all_antih_by_practice_period.dta, clear
```

```
>mmerge practice period using imported_gp_addresses_dataset.dta, type(1:1)  
missing(nomatch) unmatched(master)
```

```
>save imported_gp_addresses_coords_dataset.dta, replace
```

```
*****
```

```
*Test to see if it works - Get the basic stats
```

```
>use imported_gp_addresses_coords_dataset.dta, clear
```

```
>summarize imported_gp_addresses_dataset.dta
```

```
*Create new variable for month (shows MM instead of YYYYMM)
```

```
>gen month=substr(string(period),5,2)
```

```
*****
```

```
*import and convert vegetation tables
```

```
>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Alder_Tree.csv,  
clear
```

```
>save alder_tree_stata.dta, replace
```

```
>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Ash_Tree.csv,  
clear
```

```
>save ash_tree_stata.dta, replace
```

```
>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Birch_Tree.csv,  
clear
```

```
>save birch_tree_stata.dta, replace
```

```
>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Hazel_Tree.csv,  
clear
```

```
>save hazel_tree_stata.dta, replace
```

```

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Oak_Tree.csv,
clear

>save oak_tree_stata.dta, replace

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Pine_Tree.csv,
clear

>save pine_tree_stata.dta, replace

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Willow_Tree.csv,
clear

>save willow_tree_stata.dta, replace

*merge the datasets together based on their species, creating one dataset with all the
information

>use alder_tree_stata.dta, clear

>mmerge zonalsoa using ash_tree_stata.dta, type (1:1) missing(nomatch) unmatched(master)

*Save the new dataset

>save alder_ash_trees.dta, replace

*Repeat until all trees have been merged

>use alder_ash_trees.dta, clear

>mmerge zonalsoa using birch_tree_stata.dta, type (1:1) missing(nomatch)
unmatched(master)

>save alder_ash_birch_trees.dta, replace

>use alder_ash_birch_trees.dta, clear

>mmerge zonalsoa using hazel_tree_stata.dta, type (1:1) missing(nomatch)
unmatched(master)

>save alder_ash_birch_hazel_trees.dta, replace

>use alder_ash_birch_hazel_trees.dta, clear

>mmerge zonalsoa using oak_tree_stata.dta, type (1:1) missing(nomatch) unmatched(master)

>save alder_ash_birch_hazel_oak_trees.dta, replace

```



```

>use alder_ash_birch_hazel_oak_trees.dta, clear

>mmerge zonelsoa using pine_tree_stata.dta, type (1:1) missing(nomatch) unmatched(master)

>save alder_ash_birch_hazel_oak_pine_trees.dta, replace

>use alder_ash_birch_hazel_oak_pine_trees.dta, clear

>mmerge zonelsoa using willow_tree_stata.dta, type (1:1) missing(nomatch)
unmatched(master)

>save all_trees.dta, replace

*****

*Repeat for grasses and weeds

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Dock_Weed.csv,
clear

>save dock_weed.dta, replace

>import delimited
E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Mugwort_Weed.csv, clear

>save mugwort_weed.dta, replace

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Nettle_Weed.csv,
clear

>save nettle_weed.dta, replace

>import delimited
E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Plantain_Weed.csv, clear

>save plantain_weed.dta, replace

>import delimited
E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Ragweed_Weed.csv, clear

>save ragweed_weed.dta, replace

>use dock_weed.dta, clear

>mmerge zonelsoa using mugwort_weed.dta, type (1:1) missing(nomatch) unmatched(master)

>save dock_mugwort.dta, replace

```

```

>use dock_mugwort.dta

>mmerge zonetsoa using nettle_weed.dta, type (1:1) missing(nomatch) unmatched(master)

>save dock_mugwort_nettle.dta, replace

>use dock_mugwort_nettle.dta

>mmerge zonetsoa using plantain_weed.dta, type (1:1) missing(nomatch) unmatched(master)

>save dock_mugwort_nettle_plantain.dta, replace

>use dock_mugwort_nettle_plantain.dta, clear

>mmerge zonetsoa using ragweed_weed.dta, type (1:1) missing(nomatch) unmatched(master)

>save all_weeds.dta, replace

>import delimited E:\Projects\Pollen_LT\gp_prescribing\Zonal_Stats_Tables\Grass.csv, clear

>save Grass.dta, replace

```

\*\*\*\*\*

\*Export files to comma delimited .csv file for use in Arc GIS

```

>use all_trees.dta

>outsheet objectid zonetsoa sum_alder sum_ash sum_birch sum_hazel sum_oak sum_pine
sum_willow using all_tree_data.csv , comma

>type all_trees.csv

>use all_weeds.dta

>outsheet objectid zonetsoa mean_dock mean_mugwort mean_nettle mean_plantain
mean_ragweed using all_weed_data.csv , comma

>type all_weeds.csv

```

\*\*\*\*\*

\*Join the datasets together spatially using Arc GIS (assigning vegetation data etc to LSOAs/  
raduis')

\*In Arc GIS, Add in the initial filtered dataset, and join the vegetation data to it

\*Use the equation  $\pi \times r^2$  to create a buffer based on the radius around each GP, in which we can measure ///

\*the greenspace people within the catchment area are exposed to

\*Use the LSOA boundaries file supplied and spatially join the two to create a visual display of the data

\* Use GME to spatially join vegetation data to the polygons and points already in the data, creating AWM and AWS ///

\* in Arc GIS

\*\*\*\*\*

\*Import AWS and AWM Files

>import delimited

"E:\Projects\Pollen\_LT\gp\_prescribing\gp\_prate\_\_addressesxy\_aws\_aws\_poll.csv", clear

>save "E:\Projects\Pollen\_LT\gp\_prescribing\gp\_prate\_\_addressesxy\_aws\_aws\_poll.dta", replace

\*merge IMD and Urban/Rural into gp prescription data

>use "E:\Projects\Pollen\_LT\gp\_prescribing\gp\_prate\_\_addressesxy\_aws\_aws\_poll.dta", clear

>mmerge practice using "E:\Projects\Pollen\_LT\gp\_prescribing\gp\_info\IMD\_2010.dta",  
umatch(Practice) type(n:1) missing(nomatch) unmatched(master)

>mmerge pc\_nospace using

"E:\Projects\Pollen\_LT\gp\_prescribing\gp\_info\GP\_postcode\_E\_N\_U\_R.dta", ///

>umatch(Postcode) type(n:1) missing(nomatch) unmatched(master) ukeep(Postcode lsoa11cd  
ruc11cd urban2011\_3)

>save

"E:\Projects\Pollen\_LT\gp\_prescribing\gp\_prate\_\_addressesxy\_aws\_aws\_poll\_imd.dta",  
replace

\*\*\*\*\*

\*Import air pollution data

```

>use "E:\Projects\Pollen_LT\gp_prescribing\envdata\air_poll\pollutants_revised.dta" , clear

>save "air_poll.dta", replace

>outsheet Ozone Isoa Isoa01nm NO2 PM10 SO2 using air_pollution.csv , comma

>import delimited "E:\Projects\Pollen_LT\gp_prescribing\envdata\air_poll\air_pollution.csv",
clear

>save air_poll_actual.dta, replace

*merge air_poll into gp prescription data

>use
"E:\Projects\Pollen_LT\gp_prescribing\gp_prate_addressesxy_awm_aws_air_poll_imd.dta",
clear

>mmerge Isoa11cd using "E:\Projects\Pollen_LT\gp_prescribing\air_poll_actual.dta", ///
umatch(Isoa) type(n:1) missing(nomatch) unmatched(master) ukeep(Isoa ozone no2 pm10
so2)

>save gp_prate_addressesxy_awm_imd_airpoll.dta, replace

***Air poll_data joined in Arc GIS

*****

*import database into Arc GIS to join to existing database, removing all other duplicate
vvariables

*Create density values using the field calculator in ArcGIS, export as a .csv file using the
equation:

*(AWS/Shape_Area) = density in m squared

*Export the now-almost complete dataset as a Shapefile

*****

*Import .csv file containing densities exported from ArcGIS

>import delimited
"E:\Projects\Pollen_LT\gp_prescribing\envdata\air_poll\Analysis_densities.csv", clear

>save "Dataset_bar_ages.dta", replace

```

```

*merge in ages for GPs

>use Dataset_bar_ages.dta, clear

>mmerge practice using
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_pop_ages_2010.dta", umatch(PracCode)
type(n:1) missing(nomatch) unmatched(master)

>save analysis_dataset.dta, replace

*****

>use analysis_dataset.dta, clear

>mmerge practice using "E:\Projects\Pollen_LT\gp_prescribing\gp_info\IMD_2010.dta",
umatch(Practice) type(n:1) missing(nomatch) unmatched(master)

>mmerge pc_nospace using
"E:\Projects\Pollen_LT\gp_prescribing\gp_info\GP_postcode_E_N_U_R.dta", ///

>umatch(Postcode) type(n:1) missing(nomatch) unmatched(master) ukeep(Postcode lsoa11cd
ruc11cd urban2011_3)

>save analysis_finished_datasheet.dta, replace

*****

***Analysis Do-File***

>cd "E:\Projects\Pollen_LT\gp_prescribing\Analysis"

>use "E:\Projects\Pollen_LT\gp_prescribing\Analysis\analysis_finished_datasheet.dta", clear

*Generate month variable

>gen month=substr(string(period),5,2)

>gen month_n=real(month)

*****

*log descriptive statistics for table 1

>log using descriptive_results.txt, replace

>tabstat prate_all, by(month_n) s(mean sd n min max)

>tabstat prate_antih, by(month_n) s(mean sd n min max)

```

```

>tabstat density_alder density_ash density_birch density_hazel density_oak density_pine
density_willow , s(mean sd n min max)

>tabstat grassawm , s(mean sd n min max)

>tabstat dockawm mugwortawm nettleawm plantainawm ragweedawm , s(mean sd n min
max)

>tabstat ozoneawm no2awm pm10awm so2awm, by(month_n) s(mean sd n min max)

>tabstat radius, s(mean sd n min max)

>tabstat urban2011_3

>tabstat imd_2010, s(mean sd n min max)

>tabstat pctpop_0_14 pctpop_65up, s(mean sd n min max) columns(statistics)

>log close

*****

***Analysis Plan***

>cd "E:\Projects\Pollen_LT\gp_prescribing\Analysis"

>use analysis_finished_datasheet.dta, clear

*Model 1, Crude Model - prescribing rate = Veg Density, focussing on main 3

>log using regression_model_1.txt, replace

>regress prate_all grassawm

>regress prate_all ragweedawm

>regress prate_all density_birch

>regress prate_antih grassawm

>regress prate_antih ragweedawm

>regress prate_antih density_birch

>regress prate_all grassawm, beta

>regress prate_all ragweedawm, beta

>regress prate_all density_birch, beta

```

>regress prate\_antih grassawm, beta

>regress prate\_antih ragweedawm, beta

>regress prate\_antih density\_birch, beta

>log close

\*Model 2, Adjusted Model

\*Prescribing rate = Veg density + Air pollutants + IMD + urban/rural + % age groups

>log using regression\_model\_2.txt, replace

>regress prate\_all grassawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all ragweedawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_birch ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih grassawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih density\_birch ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all grassawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all ragweedawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_birch ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih grassawm ozoneawm no2awm so2awm pm10awm imd\_2010

i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

```
>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up, beta
```

```
>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up, beta
```

```
>log close
```

\*Sensitivity Models

\*Model 3, Adjustive Model

\*Prescribing rate = Veg density [if month>3 & month<9] + Air pollutants + IMD + urban/rural +  
% age groups

```
>log using regression_model_3.txt, replace
```

```
>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9
```

```
>regress prate_all ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <11
```

```
>regress prate_all density_birch ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>1 & month_n <6
```

```
>regress prate_antih grassawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9
```

```
>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <11
```

```
>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>1 & month_n <6
```

```
>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta
```

```
>regress prate_all ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <11, beta
```

```
>regress prate_all density_birch ozoneawm no2awm so2awm pm10awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>1 & month_n <6, beta
```



```
>regress prate_antih grassawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta
```

```
>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <11, beta
```

```
>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>1 & month_n <6, beta
```

```
>log close
```

\*Model 4, Counterfactual Model

\*Prescribing rate = Veg density [if month<4 | month>8] + Air pollutants + IMD + urban/rural +  
% age groups

```
>log using regression_model_4.txt, replace
```

```
>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>8 | month_n <4
```

```
>regress prate_all ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>10 | month_n <4
```

```
>regress prate_all density_birch ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>5 | month_n <2
```

```
>regress prate_antih grassawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>8 | month_n <4
```

```
>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>10 | month_n <4
```

```
>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>5 | month_n <2
```

```
>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>8 | month_n <4, beta
```

```
>regress prate_all ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>10 | month_n <4, beta
```

```
>regress prate_all density_birch ozoneawm no2awm so2awm pm10awm imd_2010
```

```
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>5 | month_n <2, beta
```

```

>regress prate_antih grassawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>8 | month_n <4, beta

>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>10 | month_n <4, beta

>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>5 | month_n <2, beta

>log close

*Model 5, interaction between air pollution & PM10

*Prescribing rate = Veg density [if month= pollen season]*PM10 + Other air pollutants + IMD +
urban/rural + % age groups

>log using regression_model_5.txt, replace

*Compose quintile variable

>xtile quintile= pm10awm, n(5)

>xi: regress prate_all i.quintile*grassawm ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

*store this in designated 'a'

>est store a

*run without pm10 interaction

>regress prate_all i.quintile grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

*Store the model results in a place called 'b'

>est store b

*Carry out a likelihood ratio test to compare the results of the two models

>lrtest a b

*Run calculations to measure each quintile against prescription rates

*Stratified analysis - by PM10 quintile

```

```

>xi: regress prate_all grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==1, beta

>xi: regress prate_all grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==2, beta

>xi: regress prate_all grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==3, beta

>xi: regress prate_all grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==4, beta

>xi: regress prate_all grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==5, beta

*repeat for the rest

>xi: regress prate_all i.quintile*ragweedawm ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store a

>regress prate_all i.quintile ragweedawm ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store b

>lrtest a b

>xi: regress prate_all ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==1, beta

>xi: regress prate_all ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==2, beta

>xi: regress prate_all ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==3, beta

>xi: regress prate_all ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==4, beta

>xi: regress prate_all ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==5, beta>xi: regress
prate_all

```

```
>i.quintile*density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3 pctpop_0_14  
pctpop_65up if month_n>3 & month_n <9, beta
```

```
>est store a
```

```
>regress prate_all i.quintile density_birch ozoneawm no2awm so2awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta
```

```
>est store b
```

```
>lrtest a b
```

```
>xi: regress prate_all density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==1, beta
```

```
>xi: regress prate_all density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==2, beta
```

```
>xi: regress prate_all density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==3, beta
```

```
>xi: regress prate_all density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==4, beta
```

```
>xi: regress prate_all density_birch ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==5, beta
```

```
>xi: regress prate_antih i.quintile*grassawm ozoneawm no2awm so2awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta
```

```
>est store a
```

```
>regress prate_antih i.quintile grassawm ozoneawm no2awm so2awm imd_2010  
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta
```

```
>est store b
```

```
>lrtest a b
```

```
>xi: regress prate_antih grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==1, beta
```

```
>xi: regress prate_antih grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3  
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==2, beta
```

```

>xi: regress prate_antih grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==3, beta

>xi: regress prate_antih grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==4, beta

>xi: regress prate_antih grassawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==5, beta

>xi: regress prate_antih i.quintile*ragweedawm ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store a

>regress prate_antih i.quintile ragweedawm ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store b

>lrtest a b

>xi: regress prate_antih ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==1, beta

>xi: regress prate_antih ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==2, beta

>xi: regress prate_antih ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==3, beta

>xi: regress prate_antih ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==4, beta

>xi: regress prate_antih ragweedawm ozoneawm no2awm so2awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up if month_n>3 & month_n <9 & quintile==5, beta

>xi: regress prate_antih i.quintile*density_birch ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store a

>regress prate_antih i.quintile density_birch ozoneawm no2awm so2awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up if month_n>3 & month_n <9, beta

>est store b

```

>lrtest a b

>xi: regress prate\_antih density\_birch ozoneawm no2awm so2awm imd\_2010 i.urban2011\_3  
pctpop\_0\_14 pctpop\_65up if month\_n>3 & month\_n <9 & quintile==1, beta

>xi: regress prate\_antih density\_birch ozoneawm no2awm so2awm imd\_2010 i.urban2011\_3  
pctpop\_0\_14 pctpop\_65up if month\_n>3 & month\_n <9 & quintile==2, beta

>xi: regress prate\_antih density\_birch ozoneawm no2awm so2awm imd\_2010 i.urban2011\_3  
pctpop\_0\_14 pctpop\_65up if month\_n>3 & month\_n <9 & quintile==3, beta

>xi: regress prate\_antih density\_birch ozoneawm no2awm so2awm imd\_2010 i.urban2011\_3  
pctpop\_0\_14 pctpop\_65up if month\_n>3 & month\_n <9 & quintile==4, beta

>xi: regress prate\_antih density\_birch ozoneawm no2awm so2awm imd\_2010 i.urban2011\_3  
pctpop\_0\_14 pctpop\_65up if month\_n>3 & month\_n <9 & quintile==5, beta

>log close

\*Full Veg Data

\*Model 6, Crude Model

\*prescribing rate = Veg Density,

>log using regression\_model\_6.txt, replace

>regress prate\_all grassawm

>regress prate\_all dockawm

>regress prate\_all mugwortawm

>regress prate\_all nettleawm

>regress prate\_all plantainawm

>regress prate\_all ragweedawm

>regress prate\_all density\_alder

>regress prate\_all density\_ash

>regress prate\_all density\_birch

>regress prate\_all density\_hazel

>regress prate\_all density\_oak

>regress prate\_all density\_pine  
>regress prate\_all density\_willow  
>regress prate\_antih grassawm  
>regress prate\_antih dockawm  
>regress prate\_antih mugwortawm  
>regress prate\_antih nettleawm  
>regress prate\_antih plantainawm  
>regress prate\_antih ragweedawm  
>regress prate\_antih density\_alder  
>regress prate\_antih density\_ash  
>regress prate\_antih density\_birch  
>regress prate\_antih density\_birch  
>regress prate\_antih density\_hazel  
>regress prate\_antih density\_oak  
>regress prate\_antih density\_pine  
>regress prate\_antih density\_willow  
>regress prate\_all grassawm, beta  
>regress prate\_all dockawm, beta  
>regress prate\_all mugwortawm, beta  
>regress prate\_all nettleawm, beta  
>regress prate\_all plantainawm, beta  
>regress prate\_all ragweedawm, beta  
>regress prate\_all density\_alder, beta  
>regress prate\_all density\_ash, beta  
>regress prate\_all density\_birch, beta

```

>regress prate_all density_hazel, beta
>regress prate_all density_oak, beta
>regress prate_all density_pine, beta
>regress prate_all density_willow, beta
>regress prate_antih grassawm, beta
>regress prate_antih dockawm, beta
>regress prate_antih mugwortawm, beta
>regress prate_antih nettleawm, beta
>regress prate_antih plantainawm, beta
>regress prate_antih ragweedawm, beta
>regress prate_antih density_alder, beta
>regress prate_antih density_ash, beta
>regress prate_antih density_birch, beta
>regress prate_antih density_birch, beta
>regress prate_antih density_hazel, beta
>regress prate_antih density_oak, beta
>regress prate_antih density_pine, beta
>regress prate_antih density_willow, beta
>log close

```

\*Model 7, Adjustive Model

\*Prescribing rate = Veg density + Air pollutants + IMD + urban/rural + % age groups

```

>log using regression_model_7.txt, replace

```

```

>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010

```

```

i.urban2011_3 pctpop_0_14 pctpop_65up

```

```

>regress prate_all dockawm ozoneawm no2awm so2awm pm10awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up

```



>regress prate\_all mugwortawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all nettleawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all plantainawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all ragweedawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_alder ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_ash ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_birch ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_hazel ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_oak ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_pine ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_all density\_willow ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih grassawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih dockawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih mugwortawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

>regress prate\_antih nettleawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up

```

>regress prate_antih plantainawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_alder ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_ash ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_hazel ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_oak ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_pine ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_antih density_willow ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up

>regress prate_all grassawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_all dockawm ozoneawm no2awm so2awm pm10awm imd_2010 i.urban2011_3
pctpop_0_14 pctpop_65up, beta

>regress prate_all mugwortawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_all nettleawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_all plantainawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_all ragweedawm ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

```

>regress prate\_all density\_alder ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_ash ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_birch ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_hazel ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_oak ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_pine ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_all density\_willow ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih grassawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih dockawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih mugwortawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih nettleawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih plantainawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih ragweedawm ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih density\_alder ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

>regress prate\_antih density\_ash ozoneawm no2awm so2awm pm10awm imd\_2010  
i.urban2011\_3 pctpop\_0\_14 pctpop\_65up, beta

```
>regress prate_antih density_birch ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_antih density_hazel ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_antih density_oak ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_antih density_pine ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>regress prate_antih density_willow ozoneawm no2awm so2awm pm10awm imd_2010
i.urban2011_3 pctpop_0_14 pctpop_65up, beta

>log close
```

## 2. GME Commands

```
isectpntpoly(in="E:\Projects\Pollen_LT\gp_prescribing\Merge  
data\prate_all_by_practice_GP_addressesxy.shp",  
poly="E:\Projects\Pollen_LT\gp_prescribing\Merge data\LSOA_2001_all_vegetation.shp",  
field=c("Radius", "Shape_Leng"));
```

```
isectpolypoly(in="E:\Projects\Pollen_LT\gp_prescribing\Merge  
data\LSOA_2001_all_vegetation.shp", poly="E:\Projects\Pollen_LT\gp_prescribing\Merge  
data\LSOA_GP_Radius_Buffer.shp", field="AWM", prefix="awm");
```